

Access DB# 88325

CRFE
LB

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: W.A. WALICKA Examiner #: 78201 Date: March 6, 03
Art Unit: 1652 Phone Number 305-7270 Serial Number: 09/554,414
Mail Box and Bldg/Room Location: 10 D06 Results Format Preferred (circle): PAPER DISK E-MAIL
10 D01

If more than one search is submitted, please prioritize searches in order of need. *mej*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: RNA demethylase, therapeutic & diagnostic etc
Inventors (please provide full names): MOSHE SZYF et al

Earliest Priority Filing Date: May 11 / 1998

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search

1. SEQ ID NO: 2 *o 2 p-phi*
& AA 150-411 of SEQ ID NO: 2
2. The following nucleotide seq:
CGGCGCGC *ob*
GCGCGCGC *walicka 414*

AA
2-411

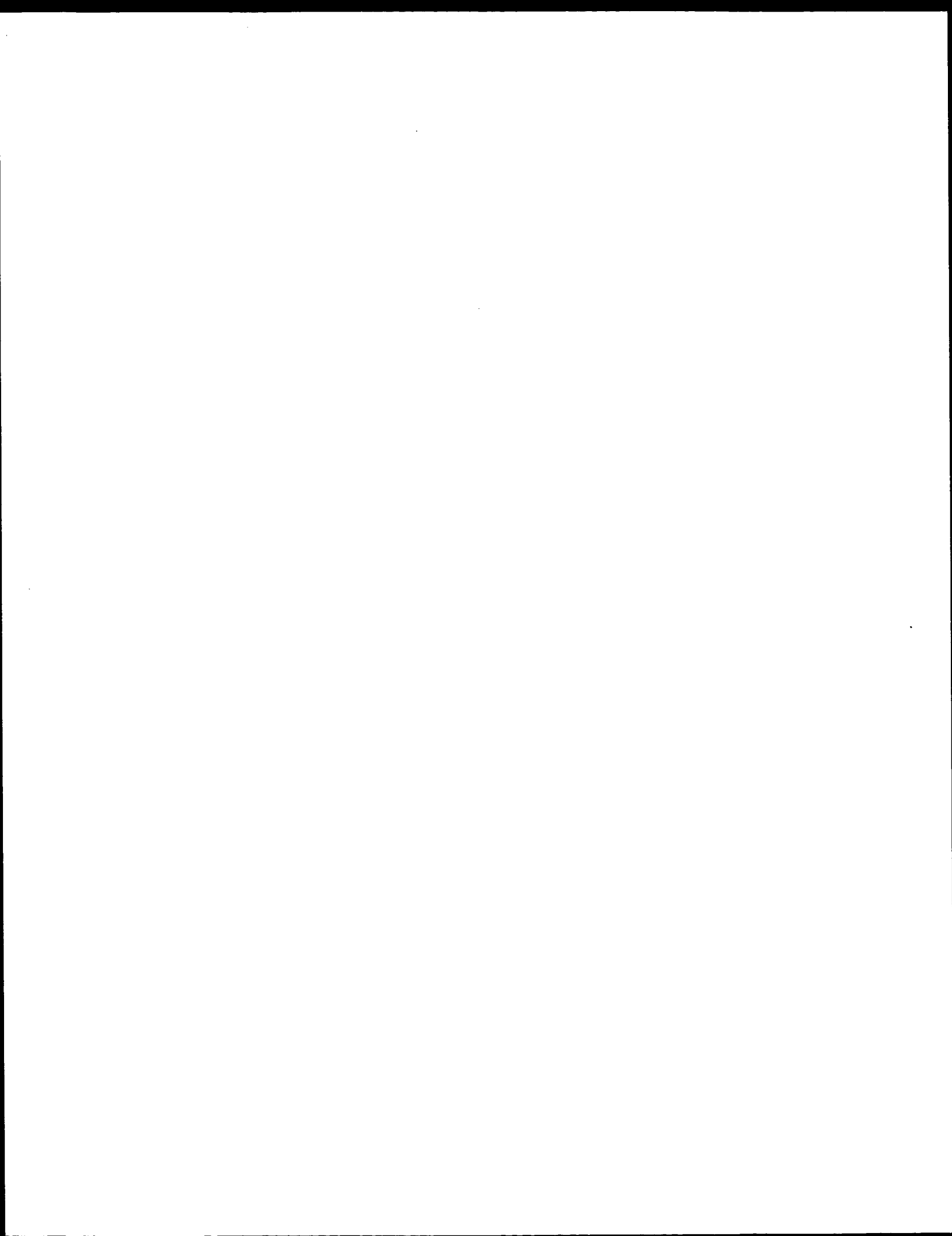
Thank you in advance.

Walicka

RECEIVED
MAR-6 2003
JACOBSON
(STIC)

STAFF USE ONLY

Searcher: <u>R. Schreiber</u>	Type of Search	Vendors and cost where applicable
Searcher Phone #: <u>308-4292</u>	Sequence (#) <u>1</u>	STN _____
Searcher Location: <u>CM1 6A03</u>	AA Sequence (#) <u>2</u>	Dialog _____
Date Searcher Picked Up: _____	Structure (#) _____	Questel/Orbit _____
Date Completed: <u>3/17</u>	Bibliographic _____	Dr. Link _____
Searcher Prep & Review Time: <u>14</u>	Litigation <u>1</u>	Lexis/Nexis _____
Clerical Prep Time: _____	Fulltext _____	Sequence Systems <u>Compu-IG</u>
Online Time: <u>17</u>	Patent Family _____	WWW/Internet _____
	Other _____	Other (specify) _____



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2003, 05:40:46 ; Search time 12.4577 Seconds
(without alignments)
618.801 Million cell updates/sec

Title: US-09-554-414b-2_COPY_150_411

Perfect score: 1344
Sequence: 1 MDCPALPQWKKKEVIRKSG.....LSRAADTEMDIEMDSDEA 262

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgcn2_6/pcdata/1/1aa/5A.COMB.pep:*
2: /cgcn2_6/pcdata/1/1aa/5B.COMB.pep:*
3: /cgcn2_6/pcdata/1/1aa/6A.COMB.pep:*
4: /cgcn2_6/pcdata/1/1aa/6B.COMB.pep:*
5: /cgcn2_6/pcdata/1/1aa/PCFUS.COMB.pep:*
6: /cgcn2_6/pcdata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	263	4	US-09-149-476-580 Sequence 580, App
2	214.5	16.0	574	4	US-09-079-431B-6 Sequence 6, App1
3	214.5	16.0	629	4	US-09-079-431B-4 Sequence 4, App1
4	214.5	16.0	630	4	US-09-079-431B-2 Sequence 2, App1
5	126	9.4	580	4	US-09-327-984A-38 Sequence 2, App1
6	97	7.2	448	3	US-08-476-509B-2 Sequence 38, App1
7	96.5	7.1	694	3	US-08-559-397A-31 Sequence 2, App1
8	95.5	7.1	1312	4	US-09-041-886-19 Sequence 31, App1
9	93	6.9	1048	3	US-09-356-952-5 Sequence 19, App1
10	92.5	6.9	459	2	US-08-870-518-1 Sequence 5, App1
11	90.5	6.7	449	1	US-08-476-008-5 Sequence 1, App1
12	90.5	6.7	449	1	US-08-476-008-7 Sequence 5, App1
13	90.5	6.7	449	1	US-08-306-063-5 Sequence 7, App1
14	90.5	6.7	449	1	US-08-306-063-7 Sequence 5, App1
15	90.5	6.7	449	1	US-08-833-485-5 Sequence 7, App1
16	90.5	6.7	449	1	US-08-833-485-7 Sequence 7, App1
17	90.5	6.7	449	4	US-09-137-440-5 Sequence 7, App1
18	90.5	6.7	449	4	US-09-137-440-7 Sequence 5, App1
19	90.5	6.7	449	5	PCT-US91-06148A-5 Sequence 7, App1
20	90.5	6.7	449	5	PCT-US91-06148A-7 Sequence 5, App1
21	88.5	6.6	1070	4	US-09-091-042A-2 Sequence 7, App1
22	88	6.5	465	5	PCT-US93-08386-5 Sequence 2, App1
23	88	6.5	466	3	US-08-348-518C-2 Sequence 2, App1
24	88	6.5	507	5	PCT-US93-08386-8 Sequence 8, App1
25	88	6.5	10182	4	US-09-134-001C-3159 Sequence 3159, Ap
26	87	6.5	861	1	US-08-484-105-18 Sequence 18, App1
27	87	6.5	861	1	US-08-484-106-18 Sequence 18, App1

28	87	6.5	2285	4	US-09-308-375-2
29	86.5	6.4	459	2	US-08-870-518-2
30	86.5	6.4	544	3	US-08-559-397A-30
31	85.5	6.4	612	1	US-08-344-695-2
32	85.5	6.4	1404	1	US-08-801-308-1
33	85.5	6.4	1792	2	US-08-962-284-4
34	84	6.2	454	3	US-08-348-518C-4
35	84	6.2	454	3	US-08-476-509B-4
36	82.5	6.1	578	4	US-09-066-046-6
37	82.5	6.1	578	4	US-08-975-762-50
38	82.5	6.1	578	4	US-09-295-028-50
39	82.5	6.1	578	4	US-09-106-582-50
40	82.5	6.1	919	2	US-08-588-983-9
41	82.5	6.1	919	2	US-08-588-976-9
42	81.5	6.1	427	2	US-09-196-857-2
43	81.5	6.1	801	1	US-08-725-012-2
44	81.5	6.1	3696	4	US-09-134-001C-5080
45	80.5	6.0	680	3	US-08-947-965-77

ALIGNMENTS

RESULT 1
US-09-149-476-580
Sequence 580, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149, 476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040, 162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038, 621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047, 600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 584
EARLIER FILING DATE: 1997-05-23

Sequence 2, App1
Sequence 2, App1
Sequence 30, App1
Sequence 2, App1
Sequence 1, App1
Sequence 4, App1
Sequence 4, App1
Sequence 4, App1
Sequence 6, App1
Sequence 50, App1
Sequence 50, App1
Sequence 9, App1
Sequence 9, App1
Sequence 2, App1
Sequence 2, App1
Sequence 5080, Ap
Sequence 77, App1

1	EARLIER APPLICATION NUMBER: 60/044,500
2	EARLIER FILING DATE: 1997-05-23
3	EARLIER APPLICATION NUMBER: 60/047,567
4	EARLIER FILING DATE: 1997-05-23
5	EARLIER APPLICATION NUMBER: 60/047,492
6	EARLIER FILING DATE: 1997-05-23
7	EARLIER APPLICATION NUMBER: 60/047,598
8	EARLIER FILING DATE: 1997-05-23
9	EARLIER APPLICATION NUMBER: 60/047,613
10	EARLIER FILING DATE: 1997-05-23
11	EARLIER APPLICATION NUMBER: 60/047,582
12	EARLIER FILING DATE: 1997-05-23
13	EARLIER APPLICATION NUMBER: 60/047,596
14	EARLIER FILING DATE: 1997-05-23
15	EARLIER APPLICATION NUMBER: 60/047,612
16	EARLIER FILING DATE: 1997-05-23
17	EARLIER APPLICATION NUMBER: 60/047,632
18	EARLIER FILING DATE: 1997-05-23
19	EARLIER APPLICATION NUMBER: 60/047,601
20	EARLIER FILING DATE: 1997-05-23
21	EARLIER APPLICATION NUMBER: 60/043,580
22	EARLIER FILING DATE: 1997-04-11
23	EARLIER APPLICATION NUMBER: 60/043,568
24	EARLIER FILING DATE: 1997-04-11
25	EARLIER APPLICATION NUMBER: 60/043,314
26	EARLIER FILING DATE: 1997-04-11
27	EARLIER APPLICATION NUMBER: 60/043,569
28	EARLIER FILING DATE: 1997-04-11
29	EARLIER APPLICATION NUMBER: 60/043,311
30	EARLIER FILING DATE: 1997-04-11
31	EARLIER APPLICATION NUMBER: 60/043,671
32	EARLIER FILING DATE: 1997-04-11
33	EARLIER APPLICATION NUMBER: 60/043,674
34	EARLIER FILING DATE: 1997-04-11
35	EARLIER APPLICATION NUMBER: 60/043,669
36	EARLIER FILING DATE: 1997-04-11
37	EARLIER APPLICATION NUMBER: 60/043,312
38	EARLIER FILING DATE: 1997-04-11
39	EARLIER APPLICATION NUMBER: 60/043,313
40	EARLIER FILING DATE: 1997-04-11
41	EARLIER APPLICATION NUMBER: 60/043,672
42	EARLIER FILING DATE: 1997-04-11
43	EARLIER APPLICATION NUMBER: 60/043,315
44	EARLIER FILING DATE: 1997-04-11
45	EARLIER APPLICATION NUMBER: 60/048,974
46	EARLIER FILING DATE: 1997-06-06
47	EARLIER APPLICATION NUMBER: 60/056,886
48	EARLIER FILING DATE: 1997-08-22
49	EARLIER APPLICATION NUMBER: 60/056,877
50	EARLIER FILING DATE: 1997-08-22
51	EARLIER APPLICATION NUMBER: 60/056,889
52	EARLIER FILING DATE: 1997-08-22
53	EARLIER APPLICATION NUMBER: 60/056,893
54	EARLIER FILING DATE: 1997-08-22
55	EARLIER APPLICATION NUMBER: 60/056,630
56	EARLIER FILING DATE: 1997-08-22
57	EARLIER APPLICATION NUMBER: 60/056,878
58	EARLIER FILING DATE: 1997-08-22
59	EARLIER APPLICATION NUMBER: 60/056,662
60	EARLIER FILING DATE: 1997-08-22
61	EARLIER APPLICATION NUMBER: 60/056,872
62	EARLIER FILING DATE: 1997-08-22
63	EARLIER APPLICATION NUMBER: 60/056,882
64	EARLIER FILING DATE: 1997-08-22
65	EARLIER APPLICATION NUMBER: 60/056,637
66	EARLIER FILING DATE: 1997-08-22
67	EARLIER APPLICATION NUMBER: 60/056,903
68	EARLIER FILING DATE: 1997-08-22
69	EARLIER APPLICATION NUMBER: 60/056,888
70	EARLIER FILING DATE: 1997-08-22
71	EARLIER APPLICATION NUMBER: 60/056,879
72	EARLIER FILING DATE: 1997-08-22
73	EARLIER APPLICATION NUMBER: 60/056,880

1	EARLIER FILING DATE: 1997-08-22	60/056, 894
2	EARLIER APPLICATION NUMBER: 60/056, 894	
3	EARLIER FILING DATE: 1997-08-22	
4	EARLIER APPLICATION NUMBER: 60/056, 911	
5	EARLIER FILING DATE: 1997-08-22	
6	EARLIER APPLICATION NUMBER: 60/056, 636	
7	EARLIER FILING DATE: 1997-08-22	
8	EARLIER APPLICATION NUMBER: 60/056, 874	
9	EARLIER FILING DATE: 1997-08-22	
10	EARLIER APPLICATION NUMBER: 60/056, 910	
11	EARLIER FILING DATE: 1997-08-22	
12	EARLIER APPLICATION NUMBER: 60/056, 864	
13	EARLIER FILING DATE: 1997-08-22	
14	EARLIER APPLICATION NUMBER: 60/056, 631	
15	EARLIER FILING DATE: 1997-08-22	
16	EARLIER APPLICATION NUMBER: 60/056, 845	
17	EARLIER FILING DATE: 1997-08-22	
18	EARLIER APPLICATION NUMBER: 60/056, 892	
19	EARLIER FILING DATE: 1997-08-22	
20	EARLIER APPLICATION NUMBER: 60/057, 761	
21	EARLIER FILING DATE: 1997-08-22	
22	EARLIER APPLICATION NUMBER: 60/047, 595	
23	EARLIER FILING DATE: 1997-05-23	
24	EARLIER APPLICATION NUMBER: 60/047, 599	
25	EARLIER FILING DATE: 1997-05-23	
26	EARLIER APPLICATION NUMBER: 60/047, 588	
27	EARLIER FILING DATE: 1997-05-23	
28	EARLIER APPLICATION NUMBER: 60/047, 585	
29	EARLIER FILING DATE: 1997-05-23	
30	EARLIER APPLICATION NUMBER: 60/047, 586	
31	EARLIER FILING DATE: 1997-05-23	
32	EARLIER APPLICATION NUMBER: 60/047, 593	
33	EARLIER FILING DATE: 1997-05-23	
34	EARLIER APPLICATION NUMBER: 60/047, 614	
35	EARLIER FILING DATE: 1997-05-23	
36	EARLIER APPLICATION NUMBER: 60/043, 578	
37	EARLIER FILING DATE: 1997-04-11	
38	EARLIER APPLICATION NUMBER: 60/043, 576	
39	EARLIER FILING DATE: 1997-04-11	
40	EARLIER APPLICATION NUMBER: 60/047, 501	
41	EARLIER FILING DATE: 1997-05-23	
42	EARLIER APPLICATION NUMBER: 60/043, 670	
43	EARLIER FILING DATE: 1997-04-11	
44	EARLIER APPLICATION NUMBER: 60/056, 632	
45	EARLIER FILING DATE: 1997-08-22	
46	EARLIER APPLICATION NUMBER: 60/056, 664	
47	EARLIER FILING DATE: 1997-08-22	
48	EARLIER APPLICATION NUMBER: 60/056, 876	
49	EARLIER FILING DATE: 1997-08-22	
50	EARLIER APPLICATION NUMBER: 60/056, 881	
51	EARLIER FILING DATE: 1997-08-22	
52	EARLIER APPLICATION NUMBER: 60/056, 909	
53	EARLIER FILING DATE: 1997-08-22	
54	EARLIER APPLICATION NUMBER: 60/056, 875	
55	EARLIER FILING DATE: 1997-08-22	
56	EARLIER APPLICATION NUMBER: 60/056, 862	
57	EARLIER FILING DATE: 1997-08-22	
58	EARLIER APPLICATION NUMBER: 60/056, 887	
59	EARLIER FILING DATE: 1997-08-22	
60	EARLIER APPLICATION NUMBER: 60/056, 908	
61	EARLIER FILING DATE: 1997-08-22	
62	EARLIER APPLICATION NUMBER: 60/048, 964	
63	EARLIER FILING DATE: 1997-06-06	
64	EARLIER APPLICATION NUMBER: 60/057, 650	
65	EARLIER FILING DATE: 1997-09-05	
66	EARLIER APPLICATION NUMBER: 60/056, 884	
67	EARLIER FILING DATE: 1997-08-22	

EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 100.0%; Score 1344; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 7,6e-123;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCPALPGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDLSFDFR 60
DB 1 MDCPALPGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDLSFDFR 60
QY 61 TCKMPSKLOKKNORLNDPLNOKNGKPDNTLPIROTASTFKOPVTKVTHNPSNKYKS 120
DB 61 TCKMPSKLOKKNORLNDPLNOKNGKPDNTLPIROTASTFKOPVTKVTHNPSNKYKS 120
QY 121 DPORMNEOPROLFEWKRLOGLSASDVTEQIITMELPKGLOGVGSNDTELLSAVASAL 180
DB 121 DPORMNEOPROLFEWKRLOGLSASDVTEQIITMELPKGLOGVGSNDTELLSAVASAL 180
QY 181 HTSSAPITGQSAVAEKNPAVNLNTSOPLCRAFIYTDIEDIRKQERVOQVRKKLEALMA 240
DB 181 HTSSAPITGQSAVAEKNPAVNLNTSOPLCRAFIYTDIEDIRKQERVOQVRKKLEALMA 240
QY 241 DILSRAADTEEMDIEMDSDEA 262
DB 241 DILSRAADTEEMDIEMDSDEA 262

RESULT 2
US-09-079-431B-6
Sequence 6, Application US/09079431B
Patent No. 6326484
GENERAL INFORMATION:
APPLICANT: Gage, Fred
APPLICANT: Ueda, Tetsuya
TITLE OF INVENTION: TRANSCRIPTION FACTOR REGULATING FGF-2
FILE REFERENCE: SALKINS.015A
CURRENT APPLICATION NUMBER: US/09/079,431B
CURRENT FILING DATE: 1998-05-14
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 574
TYPE: PRF
ORGANISM: Homo sapiens
US-09-079-431B-6

Query Match 16.0%; Score 214.5; DB 4; Length 574;
Best Local Similarity 45.7%; Pred. No. 1.7e-12;
Matches 42; Conservative 13; Mismatches 30; Indels 7; Gaps 1;

QY 1 MDCPALPGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDLSFDFR 60
DB 6 LDCPALPGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDLSFDFR 60
QY 61 TG-----KMPKSKLOKKNORLNDPLNOK 85
DB 61 TG-----KMPKSKLOKKNORLNDPLNOK 85
QY 66 GILCYPAKPAHVAVASKKRKRPRAKTRK 97
DB 66 GILCYPAKPAHVAVASKKRKRPRAKTRK 97

RESULT 3
US-09-079-431B-4
Sequence 4, Application US/09079431B
Patent No. 6326484
GENERAL INFORMATION:
APPLICANT: Gage, Fred
APPLICANT: Ueda, Tetsuya
TITLE OF INVENTION: TRANSCRIPTION FACTOR REGULATING FGF-2

TITLE OF INVENTION: AND VARIANTS THEREOF
FILE REFERENCE: SALKINS.015A
CURRENT APPLICATION NUMBER: US/09/079,431B
CURRENT FILING DATE: 1998-05-14
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 629
TYPE: PRF
ORGANISM: Homo sapiens
US-09-079-431B-4

Query Match 16.0%; Score 214.5; DB 4; Length 629;
Best Local Similarity 45.7%; Pred. No. 1.9e-12;
Matches 42; Conservative 13; Mismatches 30; Indels 7; Gaps 1;

QY 1 MDCPALPGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDLSFDFR 60
DB 6 LDCPALPGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDLSFDFR 60
QY 61 TG-----KMPKSKLOKKNORLNDPLNOK 85
DB 61 TG-----KMPKSKLOKKNORLNDPLNOK 85
QY 66 GILCYPAKPAHVAVASKKRKRPRAKTRK 97
DB 66 GILCYPAKPAHVAVASKKRKRPRAKTRK 97

RESULT 4
US-09-079-431B-2
Sequence 2, Application US/09079431B
Patent No. 6326484
GENERAL INFORMATION:
APPLICANT: Gage, Fred
APPLICANT: Ueda, Tetsuya
TITLE OF INVENTION: TRANSCRIPTION FACTOR REGULATING FGF-2
FILE REFERENCE: SALKINS.015A
CURRENT APPLICATION NUMBER: US/09/079,431B
CURRENT FILING DATE: 1998-05-14
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 630
TYPE: PRF
ORGANISM: Homo sapiens
US-09-079-431B-2

Query Match 16.0%; Score 214.5; DB 4; Length 630;
Best Local Similarity 45.7%; Pred. No. 1.9e-12;
Matches 42; Conservative 13; Mismatches 30; Indels 7; Gaps 1;

QY 1 MDCPALPGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDLSFDFR 60
DB 6 LDCPALPGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDLSFDFR 60
QY 61 TG-----KMPKSKLOKKNORLNDPLNOK 85
DB 61 TG-----KMPKSKLOKKNORLNDPLNOK 85
QY 66 GILCYPAKPAHVAVASKKRKRPRAKTRK 97
DB 66 GILCYPAKPAHVAVASKKRKRPRAKTRK 97

RESULT 5
US-09-327-984A-38
Sequence 38, Application US/09327984A
Patent No. 6368594
GENERAL INFORMATION:
APPLICANT: Doetsch, Paul W.
APPLICANT: Kaur, Balveen
APPLICANT: Avery, Angela M.
TITLE OF INVENTION: Broad Specificity DNA Damage Endonuclease
FILE REFERENCE: 25-98
CURRENT APPLICATION NUMBER: US/09/327,984A
CURRENT FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: US 60/088,521
PRIOR FILING DATE: 1998-06-08
PRIOR APPLICATION NUMBER: US 60/134,752

PRIOR FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 38
LENGTH: 580
TYPE: PRT
ORGANISM: Homo sapiens
US-09-327-984A-38

Query Match 9.4%; Score 126; DB 4; Length 580;
Best Local Similarity 22.3%; Pred. No. 0.00068;
Matches 72; Conservative 44; Mismatches 109; Indels 98; Gaps 13;

QY 5 ALPGMKKEEVIRKSGLSAGSDVYFSPSGKFRSKPOLARYLGNVDSLSPDFR 60
DB 85 SVPCGMEVYKQRLFGKTAGRDVYFISPOGLKFRKSSLANYLKKNETSLKPEDPFT 144
QY 61 T-----GKMPSKLQKNKQRLNDPLNQN-----KGRPDL----- 90
DB 145 VLSKRGISRYKDCSMALTSILOQNSNMNLTNRSKCKKDYFMPSSSELOESRGL 204
QY 91 -----NTTLPPIROTASI-----FKOPYTKVT---NHPSNKKSDPQR-----MNEQPR 130
DB 205 SMTSTHLILKEDGVDDVNFKVRKPKGKYTIILKIPITKTKGCKRSCSGFVQSDSKR 264
QY 131 QLFWEKRLQGLSASDVTEQIITKMLPKGLGVGPGSNDLTLLSAVASALHTSSAPITGQ 190
DB 265 EGVCKR-----ADAESEFVAQKSQLDRTVCSIDAGACGFTL-----SVTSE 305
QY 191 VSAAVEK-----NPAVWLNTSQPLCKAFIYTDIDIRKOE-----RVQOV 230
DB 306 ENSLVKKKERSLSSGNSNCSSEQKTSIGINKFCASAKDSEHNKEYEDTLESEIEIGKYEVV 365
QY 231 RKLEBALMADILSRADTEEMD 253
DB 366 ERK--EHLHTDLKRG--EMD 383

RESULT 6
US-08-476-509B-2
Sequence 2, Application US/08476509B
Patent No. 6034212

GENERAL INFORMATION:
APPLICANT: SUDOL, MARIUS
APPLICANT: PEER, BORK
APPLICANT: HENRY, CHEN
TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTICS
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,509B
FILING DATE: 01-DEC-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-101 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-509B-2

Query Match 7.2%; Score 97; DB 3; Length 448;
Best Local Similarity 22.0%; Pred. No. 0.31;
Matches 64; Conservative 35; Mismatches 114; Indels 78; Gaps 10;

QY 2 DCPALPGMKKEEVIRKSGLSAGSDVYFSPSGKFRSKPOLARYLGNVDSLSPDFR 61
DB 168 DVP-LPPGWEAK-----TPSGQ-----RIFLNHIDQTTWQDP 200
QY 62 GKMPSKLQKNKQRLNDPLNQNKGKPDLTTPPIROTASIFKOPYTKVTNHPSNRY--- 118
DB 201 RKAMLS--QNNVTAFTSPPYQONLMSASAMNGRISGAPV-KQPPPLAPQSPGGVMG 257
QY 119 KSPQRMNEQPRQLFWKRLQGLSASDVTEQ-----IKTMELPKGLG--VGPG----- 166
DB 258 SSSNQOQOMRLQOLQWKEKRLRLKHQELLRQELALRSQLPMTMDQDGSQNPVSSPMQOE 317
QY 167 -----SNDETLLSASALHTSSAPITGVSAVEKNPAVWLNTSQPLCKAFIYTDID 219
DB 318 LRTMTNSSDPFLNSGYHSRDESTDGSLMSYSVPRTPDDLNSVDEDTGDSISQSN 377
QY 220 IRKQEEV-----QVRRKLEBALMADILS 244
DB 378 IPSHQNRPDYLEAIPGTNVNLDGTLBGDGMNIEGELMPSLOALSSDILN 428

RESULT 7
US-08-559-397A-31
Sequence 31, Application US/08559397A
Patent No. 6083713

GENERAL INFORMATION:
APPLICANT: Manly, Susan P.
APPLICANT: Kozlowski, Michael R.
APPLICANT: Neve, Rachael L.
TITLE OF INVENTION: CLONING AND EXPRESSION OF
TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/559,397A
FILING DATE: 15-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 6013-135
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 694 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-559-397A-31

Query Match
Best Local Similarity 21.6%; Pred. No. 0.65;
Matches 65; Conservative 36; Mismatches 89; Indels 111; Gaps 12;

7.2%; Score 96.5; DB 3; Length 694;

QY 6 LPEGMKEEVIRKSGLSACK-----SDVYFSPSGKKFRS-----K 41
DB 141 LPEMEKEET--SSGISKEEQOONQAVMDIVKFDVETGDEGDMETFTTTGLPES 198
QY 42 POLA-----RYLGNITVDSFDFRTGKMMPSKLGKKRKLRLNDPLNKGKEDLNTT 93
DB 199 POVSTEPSANSPKFPSTSDSHNYGSRGT-----TPMNNHVMSTPLTND 242
QY 94 -----LPIR-----QTASIFKQPV-----TKVTNHPNKKVKSDFORMN 126
DB 243 SSSANGKFTSPRPAKPPSSASASAPIIKSPVANSANVPLKQIHATPTTPRTSPNR-- 300
QY 127 EDPRLFWERKLGSLASDVTEQIITMELPKLGCGVGSNDELTLASVASALHTSSAP 186
DB 301 -----SSISRNATLKKKEQPLPIPTKSTSPITSTATP--- 336
QY 187 ITGQVSAVEKNPA---VWLNTSOPLCRAFIYTDIEDIRKQERVOQVRKKLEALMADL 243
DB 337 ---QVVASPKPAQETVTPPTSPKPAQANSLSEKLEKREER--ERRKKQYAKLNETIC 391

QY 244 S 244
DB 392 S 392

RESULT 8
US-09-041-886-19
Sequence 19, Application US/09041886
Patent No. 6235872
GENERAL INFORMATION:
APPLICANT: Bredesen, Dale E.
APPLICANT: Rabizadeh, Sharoz
TITLE OF INVENTION: Proapoptotic Peptides, Dependence
TITLE OF INVENTION: Polypeptides and Methods of Use
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041.886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1312 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-041-886-19

Query Match
Best Local Similarity 23.5%; Pred. No. 2.1;
Matches 43; Conservative 38; Mismatches 65; Indels 37; Gaps 9;

7.1%; Score 95.5; DB 4; Length 1312;

QY 67 SKLGKKNRLNDPL--NQNKKGPDNLNTLPIRQVASTFKQVTVTNHPSKVKVSDPQRM 125
DB 748 SRLDQRO---NSPAGNENIKP--NETSP-----SFSKAEKKGISPVSEIRKQ 792
QY 126 NEOPRLFWERKLGSLASDVTEQI-----KTMELPKLGCGVGSNDELTLASVAS 178
DB 793 IDDLKKEFNDRLOPSSISEMDQLKNRGEKSRDLIK--DKIEPSANDFTIENSSN 850
QY 179 ALHTSSADITQVSAV-----EKNPAV---WLNTSOPLCRAFIYTDIEDIRKQERVOQ 229
DB 851 CTSGSSKKNPSISPSILNTEHRRGPEVTSQGVOTSPACKQ---BKDKKEKKDAEQ 907
QY 230 VRK 232
DB 908 VRK 910

RESULT 9
US-09-356-952-5
Sequence 5, Application US/09356952
Patent No. 6117663
GENERAL INFORMATION:
APPLICANT: Boriack-Sjodin, Ann
APPLICANT: Margarit, S. M.
APPLICANT: Bor-Sogil, Dafna
APPLICANT: Cole, Philip
APPLICANT: Kuriyan, John
TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 600-1-228N
CURRENT APPLICATION NUMBER: US/09/356, 952
CURRENT FILING DATE: 1999-07-19
EARLIER APPLICATION NUMBER: 60/093, 631
EARLIER FILING DATE: 1998-07-21
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 5
LENGTH: 1048
TYPE: PPT
ORGANISM: Saccharomyces cerevisiae
US-09-356-952-5

Query Match
Best Local Similarity 21.7%; Pred. No. 2.6;
Matches 52; Conservative 42; Mismatches 90; Indels 56; Gaps 12;

6.9%; Score 93; DB 3; Length 1048;

QY 54 LSSFDFTGKMMPSKLGKQKORL-----RNDELQN-----KSKPD 89
DB 85 LSRFEINNMIFHSTLFEFEREALASQKDEPERSPLQSIIGTFQKPHLRLHFLSPN 144
QY 90 LNTLPIRQVASTFKQVTVTNHPSKVKVSDPQRMNEOPRLFWERKLGSLASDVTE 148
DB 145 ELTIIP-OLTRFEKDSFNTISWNNPFLRKLRLNQHSHDLPRQMI--KAVAGASGI--VAE 200
QY 149 QIKTMELPKLGCGVGSNDELTLASVASALHTSSAP-----ITGQVSAVEKNPA 200
DB 201 NI---DELPAKQGTSCSE-----TSHSPSAFQRRRGRTJFSNVSGSDSDVT 248
QY 201 VWLNTSOPLCRAFIYTDIEDIRKQERVOQVRKKLEALMA--DILSRAADTEEN-DIEMD 257
DB 249 IWSKRKR-----YPLNETLSLVARRKKQDLGKLKQMKISANBYLSNTANSKRLNEFMN 304

RESULT 10
US-08-870-518-1
Sequence 1, Application US/08870518

APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glycosylase tolerant
NUMBER OF SEQUENCES: 5-Enolpyruvylshikimate-3-phosphate Synthases
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,008
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6047
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-008-7

Query Match 6.7% Score 90.5; DB 1; Length 449;
Best Local Similarity 22.0%; Pred. No. 1.3; Mismatches 123; Indels 41; Gaps 9;
Matches 57; Conservative 38; Mismatches 123; Indels 41; Gaps 9;
DB 16 IRKSLGAGSDVYFSSGKKFRSKPOLARYLGN-----TVDL-SSFDFTGKMPMS 67
DB 70 IRKEG-----DWIINGVNGCLOPEALDFGNAGTGARLTMGLVGTDMKTSFIDDA 123
DB 68 KLOKNNKORLNDPLNOKKGRDPL--NTLPIRQTASIFKOPVTKYTNHPNSKXVSDPQRM 125
DB 124 SLKSRPMGRVNLPLREMGVVEAADGDRMPLLTGPKTANPTTYRVPMASAOVKS----- 178
DB 126 NQOPROLFWERKLOGLSASDVT---EQIITMELPKGLOGVPGSNDFTLLSAVASALHT 182
DB 179 -----AVLLAGLNPPTVYIEPVMTNRDHEKMLQGFAGDLIVETDKDGVHRIIT 229
DB 183 SSAPITGOVSAAEKPNPAWLNTSOPLCAPITVDIEDIRKOEERVOQVQRKLEALMADI 242
DB 230 GGGKLVGQ-TIDVPGDPS---STAFPLVALLVSGSDVTIRNVLNMPTR---TGLITL 281
DB 243 LSRADTEMDIEMDSGDE 261
DB 282 QEWGADIEVILNRLAGGED 300

RESULT 13
US-08-306-063-5
Sequence 5, Application US/08306063
Patent No. 5633435
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glycosylase tolerant
NUMBER OF SEQUENCES: 5-Enolpyruvylshikimate-3-phosphate Synthases
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6047
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-306-063-5

Query Match 6.7% Score 90.5; DB 1; Length 449;
Best Local Similarity 22.0%; Pred. No. 1.3; Mismatches 123; Indels 41; Gaps 9;
Matches 57; Conservative 38; Mismatches 123; Indels 41; Gaps 9;
DB 16 IRKSLGAGSDVYFSSGKKFRSKPOLARYLGN-----TVDL-SSFDFTGKMPMS 67
DB 70 IRKEG-----DWIINGVNGCLOPEALDFGNAGTGARLTMGLVGTDMKTSFIDDA 123
DB 68 KLOKNNKORLNDPLNOKKGRDPL--NTLPIRQTASIFKOPVTKYTNHPNSKXVSDPQRM 125
DB 124 SLKSRPMGRVNLPLREMGVVEAADGDRMPLLTGPKTANPTTYRVPMASAOVKS----- 178
DB 126 NQOPROLFWERKLOGLSASDVT---EQIITMELPKGLOGVPGSNDFTLLSAVASALHT 182
DB 179 -----AVLLAGLNPPTVYIEPVMTNRDHEKMLQGFAGDLIVETDKDGVHRIIT 229
DB 183 SSAPITGOVSAAEKPNPAWLNTSOPLCAPITVDIEDIRKOEERVOQVQRKLEALMADI 242
DB 230 GGGKLVGQ-TIDVPGDPS---STAFPLVALLVSGSDVTIRNVLNMPTR---TGLITL 281
DB 243 LSRADTEMDIEMDSGDE 261

Db 282 QEMGADIEVLNARLAGGED 300

RESULT 14

US-08-306-063-7

Sequence 7, Application US/08306063

Patent No. 5633435

GENERAL INFORMATION:

APPLICANT: Barry, Gerard F.

APPLICANT: Kishore, Ganesh M.

APPLICANT: Padgett, Stephen R.

APPLICANT: Stallings, William C.

TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESS: Dennis R. Hoerner, Jr., Monsanto Co. B44F

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/306,063

FILING DATE: 13-SEP-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/749,611

FILING DATE: 28-AUG-1991

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/576,537

FILING DATE: 31-AUG-1990

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914

REFERENCE/DOCKET NUMBER: 38-21(10660)A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)537-6099

TELEFAX: (314)537-6047

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 449 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-306-063-7

Query Match 6.7%; Score 90.5; DB 1; Length 449;
Best Local Similarity 22.0%; Pred. No. 1.3; Indels 41; Gaps 9;
Matches 57; Conservative 38; Mismatches 123;

QY 16 IRKSGISACKSDVYFSPGKFRSKPOLARYLGN-----TYVDL-SFDFRTGKMPS 67
Db 70 IRREG-----DWTIINGVNGCLLPALDFGNAGARLTGVLGTITDMKTSFIGDA 123
QY 68 KIQKNQRLNDPLNKNKKRPL--NTLPIRQTASTIKQPVTKVTHNPSKKVSDPORM 125
Db 124 SLKRPKMGVNLPLREMGVVEADGDGDRMPLTLIGPKTANPITVYRVMAAQVKS----- 178
QY 126 NEOPROLFEWERKLOGLSADVT---EQIITKTMELPKGLOGVSGPSNDETLISAVALHT 182
Db 179 -----AVLLAGINTPGVTTVLEPVMTNDHTEKMLQGGADLTVEYTDKDGVRHIRT 229
QY 183 SSAPITGVSAVERNPVMTNTSOPLCARIVTDEDIRKQEEVQVQRKKLEALMADI 242
Db 230 GGGKLVQGTIDVPDPS---STAFPLVAALLVSGSDVTIRNVTLMNPTFR---TGLILTL 281

QY 243 LSRADTFEEMDIEMDSDE 261
Db 282 QEMGADIEVLNARLAGGED 300

RESULT 15

US-08-833-485-5

Sequence 5, Application US/08833485

Patent No. 5804425

GENERAL INFORMATION:

APPLICANT: Barry, Gerard F.

APPLICANT: Kishore, Ganesh M.

APPLICANT: Padgett, Stephen R.

APPLICANT: Stallings, William C.

TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESS: Dennis R. Hoerner, Jr., Monsanto Co. B44F

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/833,485

FILING DATE: 07-APR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/306,063

FILING DATE: 13-SEP-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/749,611

FILING DATE: 28-AUG-1991

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/576,537

FILING DATE: 31-AUG-1990

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914

REFERENCE/DOCKET NUMBER: 38-21(15117)A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)737-6099

TELEFAX: (314)737-6047

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 449 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-833-485-5

Query Match 6.7%; Score 90.5; DB 1; Length 449;
Best Local Similarity 22.0%; Pred. No. 1.3; Indels 41; Gaps 9;
Matches 57; Conservative 38; Mismatches 123;

QY 16 IRKSGISACKSDVYFSPGKFRSKPOLARYLGN-----TYVDL-SFDFRTGKMPS 67
Db 70 IRREG-----DWTIINGVNGCLLPALDFGNAGARLTGVLGTITDMKTSFIGDA 123
QY 68 KIQKNQRLNDPLNKNKKRPL--NTLPIRQTASTIKQPVTKVTHNPSKKVSDPORM 125
Db 124 SLKRPKMGVNLPLREMGVVEADGDGDRMPLTLIGPKTANPITVYRVMAAQVKS----- 178

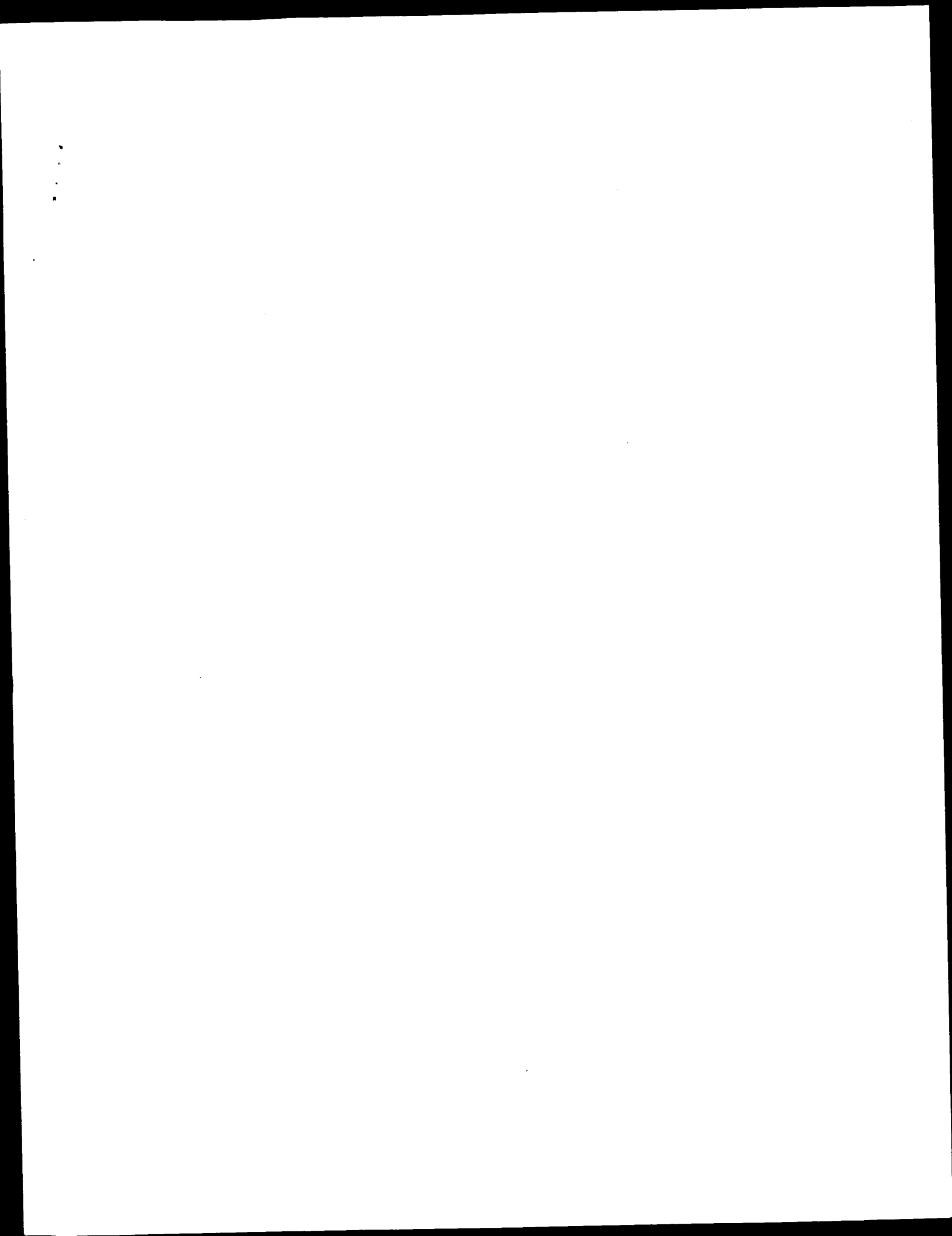
Mon Mar 17 08:44:14 2003

us-09-554-414b-2_copy_150_411.rai

Page 9

QY 126 NEQPROLFWEKRLQGLSASDVT---EQIKTMELPKGLOGVPGSNDFTLSAVASALHT 182
Db 179 -----AVLLAGLNTPGVTTVIEPVMTRDHTKMLQGFADLTVEFDKGVRHIRT 229
QY 183 SSAPITGOYSAVAKRNPAWMLNTSOPLCFAFVTDIEDIRKQERYQOVKKLEBALMADI 242
Db 230 GCGKLVGO-TIDVPGDPs---STAFPLVALLVESGSDVTIRNVLMNPTTR---TGLITTL 281
QY 243 LSRADTEMDIEMDSGDE 261
Db 282 QEWGADIEVLNARLAGED 300

Search completed: March 12, 2003, 09:15:37
Job time : 17.4577 secs



GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2003, 03:17:33 ; Search time 36.2051 Seconds
(Without alignments)
1491.072 Million cell updates/sec

Title: US-09-554-414B-2_COPY_150_411

Perfect score: 1344

Sequence: 1 MDCPALPGMKKEVIRKSG.....LSRADTEEMDIEMSGDEA 262

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_minc:*
9: sp_organella:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	411	4 Q9UBB5	Q9UBB5 homo sapien
2	1337	98.7	414	11 Q922E1	Q922E1 mus musculu
3	1050	78.1	282	13 Q9PUM9	Q9PUM9 xenopus lae
4	1047.5	77.9	303	13 Q9PUM8	Q9PUM8 xenopus lae
5	1021	76.0	291	4 Q95983	Q95983 homo sapien
6	1013	75.4	200	4 Q60535	Q60535 homo sapien
7	1012.5	75.3	285	11 Q922B8	Q922B8 mus musculu
8	451	33.6	302	4 Q9UIS8	Q9UIS8 homo sapien
9	446	33.2	249	11 Q922D9	Q922D9 mus musculu
10	436.5	32.5	314	5 Q9YHB7	Q9YHB7 mus musculu
11	409	30.4	146	4 Q95242	Q95242 dirosophila
12	384.5	28.6	226	5 Q9V424	Q9V424 homo sapien
13	339	25.2	194	4 Q8WYV6	Q8WYV6 homo sapien
14	246	18.3	186	11 Q9D2S3	Q9D2S3 mus musculu
15	246	18.3	186	11 Q9D9H3	Q9D9H3 mus musculu
16	237.5	17.7	588	11 Q9DC19	Q9DC19 mus musculu

17	230.5	17.2	636	11 Q922E2	Q922E2 mus musculu
18	214.5	16.0	277	4 Q96EC1	Q96EC1 homo sapien
19	214.5	16.0	503	4 Q9UNZ6	Q9UNZ6 homo sapien
20	214.5	16.0	549	4 Q9UNZ7	Q9UNZ7 homo sapien
21	214.5	16.0	556	4 Q15248	Q15248 homo sapien
22	214.5	16.0	586	4 Q9UNZ8	Q9UNZ8 homo sapien
23	214.5	16.0	605	4 Q9UNZ9	Q9UNZ9 homo sapien
24	214.5	16.0	605	4 Q9UIS9	Q9UIS9 homo sapien
25	177	13.2	486	6 Q95L68	Q95L68 macaca fasc
26	167	12.4	467	13 Q9YGC6	Q9YGC6 xenopus lae
27	164.5	12.2	344	13 Q42403	Q42403 gallus gall
28	135.5	10.1	554	11 Q8R3R3	Q8R3R3 mus musculu
29	130.5	9.7	554	11 Q922D7	Q922D7 mus musculu
30	126	9.4	225	10 Q9L7I1	Q9L7I1 arabidopsis
31	126	9.4	574	4 Q96F09	Q96F09 homo sapien
32	126	9.4	580	4 Q95243	Q95243 homo sapien
33	125.5	9.3	155	10 Q9F2P6	Q9F2P6 arabidopsis
34	119	8.9	1627	5 Q962D0	Q962D0 giaridia lam
35	118	8.8	186	10 Q9LYB9	Q9LYB9 arabidopsis
36	117.5	8.7	384	10 Q9X136	Q9X136 arabidopsis
37	115	8.6	270	10 Q9F2P7	Q9F2P7 arabidopsis
38	115	8.6	353	10 Q9SLR5	Q9SLR5 tritium ae
39	115	8.6	820	10 Q9STZ6	Q9STZ6 arabidopsis
40	111	8.3	2003	11 Q91Y51	Q91Y51 mus musculu
41	111	8.3	2025	11 Q9R0L6	Q9R0L6 mus musculu
42	110.5	8.2	602	13 Q8OH33	Q8OH33 brachydanio
43	109.5	8.1	182	10 Q9SNC0	Q9SNC0 arabidopsis
44	108.5	8.1	865	5 Q9N4L7	Q9N4L7 caenorhabdit
45	108.5	8.1	1972	4 Q9UIF8	Q9UIF8 homo sapien

ALIGNMENTS

RESULT 1
Q9UBB5 PRELIMINARY: PRT; 411 AA.
ID Q9UBB5
AC Q9UBB5;
DT 01-MAY-2000 (TREMBLrel. 13. Created)
DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)
DT 01-MAY-2002 (TREMBLrel. 20. Last annotation update)
DE Methyl-Cpg binding protein 2.
DE MB02.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99373255; PubMed=10441743;
RA Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;
RT "Genomic structure and chromosomal mapping of the murine and human
RL mbdl, mbd2, mbd3, and mbd4 genes.";
RM Mann. Genome 10:906-912(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9843942; PubMed=9774669;
RA Hendrich B., Bird A.;
RT "Identification and characterization of a family of mammalian methyl-
RT Cpg binding proteins.";
RM Mol. Cell. Biol. 18:6538-6547(1998).
DR EMBL; AF120993; AAD56597.1; JOINED.
DR EMBL; AF120988; AAD56597.1; JOINED.
DR EMBL; AF120989; AAD56597.1; JOINED.
DR EMBL; AF120990; AAD56597.1; JOINED.
DR EMBL; AF120991; AAD56597.1; JOINED.
DR EMBL; AF120992; AAD56597.1; JOINED.
DR EMBL; AF072242; AAC68871.1; JOINED.
DR Interpro: IPR001739; Methyl-Cpg-bind.
DR Pfam: PF01429; MBD; 1.
DR SMART: SM00391; MBD; 1.
SQ SEQUENCE 411 AA; 43254 MW; FC4E5E0CF9BA0FPA CRC64;

Query Match 100.0%; Score 1344; DB 4; Length 411;
 Best Local Similarity 100.0%; Pred. No. 7.8e-101;
 Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCPALPGWKKEEYIRKSGLSAGKSDVYFSPSGKFRSRKPOLARYLGNVTDLSSPDR 60
 DB 150 MDCPALPGWKKEEYIRKSGLSAGKSDVYFSPSGKFRSRKPOLARYLGNVTDLSSPDR 209

QY 61 TGMKMPSTLQKNNKQRLNDPLNOKKGPDLNTLPIRQTASIFKQPYTKVTHNPSKNVKS 120
 DB 210 TGMKMPSTLQKNNKQRLNDPLNOKKGPDLNTLPIRQTASIFKQPYTKVTHNPSKNVKS 269

QY 121 DPQRNNEOPROLFEWKRLOGLSASDVTEQIKTMELPKLOGVGPSSNDETLISAVASAL 180
 DB 270 DPQRNNEOPROLFEWKRLOGLSASDVTEQIKTMELPKLOGVGPSSNDETLISAVASAL 329

QY 181 HTSSAPITGOVSAAEKNPAAVWLNTPSOPLCKAFITVDEDIRKOEERVOOVRKKLEALMA 240
 DB 330 HTSSAPITGOVSAAEKNPAAVWLNTPSOPLCKAFITVDEDIRKOEERVOOVRKKLEALMA 389

QY 241 DILSRADTEEMDIEMDSGDEA 262
 DB 390 DILSRADTEEMDIEMDSGDEA 411

RESULT 2
 Q9Z2EI PRELIMINARY; PRT; 414 AA.
 AC Q9Z2EI;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Methyl-Cpg binding protein MBD2.
 GN MBD2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=9844942; PubMed=9774669;
 RA Hendrich B., Bird A.;
 RT "Identification and characterization of a family of mammalian methyl-
 Cpg binding proteins."
 RT Mol. Cell. Biol. 18:6538-6547(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=99373255; PubMed=10441743;
 RA Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;
 RT "Genomic structure and chromosomal mapping of the murine and human
 mbd1, mbd2, mbd3, and mbd4 genes."
 RT Mann. Genome 10:906-912(1999).
 DR EMBL: AF072243; AAC68872.1; -
 DR EMBL: AF120986; AAD50372.1; -
 DR EMBL: AF120983; AAD50372.1; JOINED.
 DR EMBL: AF120984; AAD50372.1; JOINED.
 DR EMBL: AF120985; AAD50372.1; JOINED.
 DR MGD: MGI:133813; Mbd2.
 DR InterPro: IPR001739; Methyl-Cpg_bind.
 DR Pfam: PF01429; MBD.1.
 DR SMART: SM00391; MBD.1.
 SQ SEQUENCE 414 AA; 43543 MW; 9601D95E347E8E53 CRC64;

Query Match 98.7%; Score 1327; DB 11; Length 414;
 Best Local Similarity 98.5%; Pred. No. 1.9e-99;
 Matches 258; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDCPALPGWKKEEYIRKSGLSAGKSDVYFSPSGKFRSRKPOLARYLGNVTDLSSPDR 60
 DB 153 MDCPALPGWKKEEYIRKSGLSAGKSDVYFSPSGKFRSRKPOLARYLGNVTDLSSPDR 212

QY 61 TGMKMPSTLQKNNKQRLNDPLNOKKGPDLNTLPIRQTASIFKQPYTKVTHNPSKNVKS 120
 DB 213 TGMKMPSTLQKNNKQRLNDPLNOKKGPDLNTLPIRQTASIFKQPYTKVTHNPSKNVKS 272

QY 121 DPQRNNEOPROLFEWKRLOGLSASDVTEQIKTMELPKLOGVGPSSNDETLISAVASAL 180
 DB 273 DPQRNNEOPROLFEWKRLOGLSASDVTEQIKTMELPKLOGVGPSSNDETLISAVASAL 332

QY 181 HTSSAPITGOVSAAEKNPAAVWLNTPSOPLCKAFITVDEDIRKOEERVOOVRKKLEALMA 240
 DB 333 HTSSAPITGOVSAAEKNPAAVWLNTPSOPLCKAFITVDEDIRKOEERVOOVRKKLEALMA 392

QY 241 DILSRADTEEMDIEMDSGDEA 262
 DB 393 DILSRADTEEMDIEMDSGDEA 414

RESULT 3
 Q9PUM9 PRELIMINARY; PRT; 282 AA.
 AC Q9PUM9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Methyl-Cpg binding protein MBD3.
 GN MBD3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OX Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wade P.A., Geggion A., Jones P.L., Ballestar E., Andry F.,
 RA Wolfe A.P.;
 RT "The M1-2 histone deacetylase complex couples DNA methylation to
 RT chromatin remodeling and histone deacetylation."
 RL Nat. Genet. 0:0-0(1999).
 DR EMBL: AF170346; AAD55389.1; -
 DR InterPro: IPR001739; Methyl-Cpg_bind.
 DR Pfam: PF01429; MBD.1.
 DR SMART: SM00391; MBD.1.
 SQ SEQUENCE 282 AA; 31692 MW; 6891EFA05B0371E9 CRC64;

Query Match 78.1%; Score 1050; DB 13; Length 282;
 Best Local Similarity 75.5%; Pred. No. 3.2e-77;
 Matches 200; Conservative 33; Mismatches 26; Indels 6; Gaps 2;

QY 2 DCPALPGWKKEEYIRKSGLSAGKSDVYFSPSGKFRSRKPOLARYLGNVTDLSSPDR 61
 DB 7 ECSSL-QGWKKEEYIRKSGLSAGKSDVYFSPSGKFRSRKPOLARYLGNVTDLSSPDR 65

QY 62 GKMMPSTLQKNNKQRLNDPLNOKKGPDLNTLPIRQTASIFKQPYTKVTHNPSKNVKS 121
 DB 66 GKMMPSTLQKNNKQRLNDPLNOKKGPDLNTLPIRQTASIFKQPYTKVTHNPSKNVKS 125

QY 122 DPQRNNEOPROLFEWKRLOGLSASDVTEQIKTMELPKLOGVGPSSNDETLISAVASAL 181
 DB 126 DPQRNNEOPROLFEWKRLOGLSASDVTEQIKTMELPKLOGVGPSSNDETLISAVASAL 185

QY 182 TSSAPITGOVSAAEKNPAAVWLNTPSOPLCKAFITVDEDIRKOEERVOOVRKKLEALMA 241
 DB 186 TSSAPITGOVSAAEKNPAAVWLNTPSOPLCKAFITVDEDIRKOEERVOOVRKKLEALMA 245

QY 242 DILSRADTEEMDIEMDSGDEA 261
 DB 246 DILSRADTEEMDIEMDSGDEA 270

RESULT 4
 Q9PUM8 PRELIMINARY; PRT; 303 AA.
 AC Q9PUM8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Methyl-Cpg binding protein MBD3.
 GN MBD3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OX Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wade P.A., Geggion A., Jones P.L., Ballestar E., Andry F.,
 RA Wolfe A.P.;
 RT "The M1-2 histone deacetylase complex couples DNA methylation to
 RT chromatin remodeling and histone deacetylation."
 RL Nat. Genet. 0:0-0(1999).
 DR EMBL: AF170346; AAD55389.1; -
 DR InterPro: IPR001739; Methyl-Cpg_bind.
 DR Pfam: PF01429; MBD.1.
 DR SMART: SM00391; MBD.1.
 SQ SEQUENCE 303 AA; 32712 MW; 6891EFA05B0371E9 CRC64;

DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Methyl-CpG binding protein MBD3 long form.
 GN MBD3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wade P.A., Gegonne A., Jones P.L., Ballestar E., Aubry F.,
 RA Wolfe A.P.;
 RT "The M1-2 histone deacetylase complex couples DNA methylation to
 RT chromatin remodeling and histone deacetylation."
 RL Nat. Genet. 0:0-0(1999)
 DR EMBL: AF170347; AAC55390.1; -
 DR InterPro: IPR001739; Methyl-CpG_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SO SEQUENCE 303 AA; 34078 MW; 864B727E50EF710B CRC64;

Query Match 77.9%; Score 1047.5; DB 13; Length 303;
 Best Local Similarity 70.5%; Pred. No. 5.6e-77;
 Matches 201; Conservative 33; Mismatches 26; Indels 25; Gaps 2;

QY 2 DCPALPGMKKEEYIRKSGLSAGSDVYFSSPSGKFRSKPOLARYLGNTVDSFDFRT 1
 DB 7 ECSPALPGMKKEEYIRKSGLSAGSDVYFSSPSGKFRSKPOLARYLGNTVDSFDFRT 66
 QY 42 POLARYLGNTVDSFDFRTGKMPKSLQKQKRLRNDPLNKGKPDNTLPIRQTAS 101
 DB 67 POLARYLGNTVDSFDFRTGKMPKSLQKQKRLRNDPLNKGKPDNTLPIRQTAS 126
 QY 102 IFKOPVTKVTHNPNKYSKSPQKRAVDQRLFEKRLQGLASADYEQITKTELKGLQ 161
 DB 127 IFKOPVTKVTHNPNKYSKSPQKRAVDQRLFEKRLQGLASADYEQITKTELKGLQ 186
 QY 162 GVGPGSNDLTLVSAVSAHTSSAPITGQVSAVENKPAWMLTSPOLCAFIPTDEDIR 221
 DB 187 GVGPGSNDLTLVSAVSAHTSSAPITGQVSAVENKPAWMLTSPOLCAFIPTDEDIR 246
 QY 222 KOEERVQVKKLEALMDILSRADTEE-----MDIEMDSGE 261
 DB 247 KOEELVQVKKLEALMDILSRADTEE-----MDIEMDSGE 291

RESULT 5

ID 095983 PRELIMINARY; PRT; 291 AA.
 AC 095983;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Methyl-CpG binding protein MBD3.
 GN MBD3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCready P.M., Skowronski E., Vismathan V.,
 RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stillgen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnea J.,
 RA Danganan L., Eiler A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Altix C., Andreise T., Tranheim M., Amico-Keller G., Coefield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
 RA Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carrano A.V.;
 RT "Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a
 RT serine protease gene cluster."

RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98449942; PubMed=9774669;
 RT Hendrich B., Bird A.;
 RT "Identification and characterization of a family of mammalian methyl-
 RT CpG binding proteins."
 RL Mol. Cell. Biol. 18:6538-6547(1998).
 DR EMBL: AC005943; AAC72104.1; -
 DR EMBL: AF072247; AAC68876.1; -
 DR InterPro: IPR001739; Methyl-CpG_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SO SEQUENCE 291 AA; 32844 MW; B62134DD1BEB636B CRC64;

Query Match 76.0%; Score 1021; DB 4; Length 291;
 Best Local Similarity 72.1%; Pred. No. 7.5e-75;
 Matches 196; Conservative 34; Mismatches 30; Indels 12; Gaps 2;

QY 2 DCPALPGMKKEEYIRKSGLSAGSDVYFSSPSGKFRSKPOLARYLGNTVDSFDFRT 61
 DB 7 ECSPALPGMKKEEYIRKSGLSAGSDVYFSSPSGKFRSKPOLARYLGNTVDSFDFRT 66
 QY 62 GKMPKSLQKQKRLRNDPLNKGKPDNTLPIRQTASIFKOPVTKVTHNPNKYSKSP 121
 DB 67 GKMPKSLQKQKRLRNDPLNKGKPDNTLPIRQTASIFKOPVTKVTHNPNKYSKSP 126
 QY 122 PORNAEQPOLFEKRLQGLASADYEQITKTELKGLQGVGPGSNDLTLVSAVSA 181
 DB 127 PORNAEQPOLFEKRLQGLASADYEQITKTELKGLQGVGPGSNDLTLVSAVSA 186
 QY 182 TSSAPITGQVSAVENKPAWMLTSPOLCAFIPTDEDIRKOEERVQVKKLEALMD 241
 DB 187 TSSAPITGQVSAVENKPAWMLTSPOLCAFIPTDEDIRKOEERVQVKKLEALMD 246
 QY 242 ILSR-----AADTE-----EMDIEMDSGE 261
 DB 247 MLHVEELARDEAPLDKACAEDEDEDEE 278

RESULT 6

ID 060535 PRELIMINARY; PRT; 200 AA.
 AC 060535;
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Antigen NY-CO-41 (Fragment).
 GN NY-CO-41.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON CANCER METASTASIS TO LUNG;
 RX MEDLINE=98272252; PubMed=9610721;
 RA Scanlan M.J., Chen Y.T., Williamson B., Gure A.O., Stockert E.,
 RA Gordon J.B., Tureci O., Sahin U., Pfreundschuh M., Old L.J.;
 RT "Characterization of human colon cancer antigens recognized by
 RT autologous antibodies."
 RL Int. J. Cancer 76:652-658(1998).
 DR EMBL: AF039701; AAC18050.1; -
 DR NON_TER
 SO SEQUENCE 200 AA; 22270 MW; A70B46B35CA55AFD CRC64;

Query Match 75.4%; Score 1013; DB 4; Length 200;
 Best Local Similarity 100.0%; Pred. No. 2e-74;
 Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 KMPKSLQKQKRLRNDPLNKGKPDNTLPIRQTASIFKOPVTKVTHNPNKYSKSP 122
 DB 1 KMPKSLQKQKRLRNDPLNKGKPDNTLPIRQTASIFKOPVTKVTHNPNKYSKSP 60

QY 123 QRMNEOPROLFWERKLOGLSASDVTEQIITKTMELPKGLGCVGSGSNDTELLSAVALHT 182
 DB 61 QRMNEOPROLFWERKLOGLSASDVTEQIITKTMELPKGLGCVGSGSNDTELLSAVALHT 120
 QY 183 SSAPITGOVSAAEKNPAVWINTSOPLCAPITVDEDIRKOEYVQVQRKLEALMADI 242
 DB 121 SSAPITGOVSAAEKNPAVWINTSOPLCAPITVDEDIRKOEYVQVQRKLEALMADI 180
 QY 243 LSRAADTEEMDENDSGDEA 262
 DB 181 LSRAADTEEMDENDSGDEA 200

RESULT 7

Q922D8 PRELIMINARY; PRT; 285 AA.
 ID Q922D8
 AC Q922D8: (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Methyl-Cpg binding protein MBD3.
 GN MBD3.
 OS Mus musculus (Mouse), and
 OS Mus musculus domesticus (western European house mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090, 10092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98449942; PubMed=9774669;
 RA Hendrich B., Bird A.,
 RT "Identification and characterization of a family of mammalian methyl-
 Cpg binding proteins."
 RL Mol. Cell. Biol. 18:6538-6547(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=99373255; PubMed=10441743;
 RA Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;
 RT "Genomic structure and chromosomal mapping of the murine and human
 Mbd1, mbd2, mbd3, and mbd4 genes."
 RL Mamm. Genome 10:906-912(1999).
 DR EMBL: AF072248; AAC68877.1; -;
 DR EMBL: AF120989; AAD48909.1; -;
 DR MGI:133812; Mbd3.
 DR InterPro: IPR001739; Methyl-Cpg_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SQ SEQUENCE 285 AA; 32168 MW; EAE57BD48463643F CRC64;

Query Match 75.3%; Score 1012.5; DB 11; Length 285;
 Best Local Similarity 75.3%; Pred. No. 3.6e-74;
 Matches 189; Conservative 35; Mismatches 24; Indels 3; Gaps 1;

QY 2 DCPALPPGKKKEVIRKSGLSAGKSDVYFFSPGKFRSKPOLARYLGNTVLDLSSFDRT 61
 DB 7 ECPALPPGKKKEVIRKSGLSAGKSDVYFFSPGKFRSKPOLARYLGNTVLDLSSFDRT 66
 QY 62 GKMMPKSLQKNKORLNDPLNONGKRPDLNTLPIROTASIFKOPVKTYNHPSNKKVSD 121
 DB 67 GKMMPKSLQKNKORLNDPLNONGKRPDLNTLPIROTASIFKOPVKTYNHPSNKKVSD 126
 QY 122 PQRMBDOPROLFWERKLOGLSASDVTEQIITKTMELPKGLGCVGSGSNDTELLSAVALH 181
 DB 127 PQRMBDOPROLFWERKLOGLSASDVTEQIITKTMELPKGLGCVGSGSNDTELLSAVALH 186
 QY 182 TSSAPITGOVSAAEKNPAVWINTSOPLCAPITVDEDIRKOEYVQVQRKLEALMADI 241
 DB 187 TSSAPITGOVSAAEKNPAVWINTSOPLCAPITVDEDIRKOEYVQVQRKLEALMADI 246
 QY 242 ILSRAADTEEM 252

DB 247 ML---AHVEL 254

RESULT 8

Q9UIS8 PRELIMINARY; PRT; 302 AA.
 ID Q9UIS8
 AC Q9UIS8: (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Testis-specific methyl-Cpg binding protein 2.
 GN MBD2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99373255; PubMed=10441743;
 RA Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;
 RT "Genomic structure and chromosomal mapping of the murine and human
 Mbd1, Mbd2, Mbd3, and Mbd4 genes."
 RL Mamm. Genome 10:906-912(1999).
 DR EMBL: AF120989; AAD56596.1; -;
 DR EMBL: AF120988; AAD56596.1; JOINED.
 DR InterPro: IPR001739; Methyl-Cpg_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SQ SEQUENCE 302 AA; 31745 MW; CCEC65D926222717 CRC64;

Query Match 33.6%; Score 451; DB 4; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.3e-28;
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCPALPPGKKKEVIRKSGLSAGKSDVYFFSPGKFRSKPOLARYLGNTVLDLSSFDRT 60
 DB 150 MDCPALPPGKKKEVIRKSGLSAGKSDVYFFSPGKFRSKPOLARYLGNTVLDLSSFDRT 209
 QY 61 TGMMPKSLQKNKORLNDPLNONGK 85
 DB 210 TGMMPKSLQKNKORLNDPLNONGK 234

RESULT 9

Q922D9 PRELIMINARY; PRT; 249 AA.
 ID Q922D9
 AC Q922D9: (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Testis specific methyl-Cpg binding protein MBD2.
 GN MBD2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6; TISSUE=TESTIS;
 RX MEDLINE=98449942; PubMed=9774669;
 RA Hendrich B., Bird A.,
 RT "Identification and characterization of a family of mammalian methyl-
 Cpg binding proteins."
 RL Mol. Cell. Biol. 18:6538-6547(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=99373255; PubMed=10441743;
 RA Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;
 RT "Genomic structure and chromosomal mapping of the murine and human
 Mbd1, mbd2, mbd3, and mbd4 genes."
 RL Mamm. Genome 10:906-912(1999).
 DR EMBL: AF072245; AAC68874.1; -;

Db	1	PGMKKEWIRKSGLSACKSDVYVEYSPSGKKFRKSPQLAYLTNTVDLSFDRPTGKMPS	60
Qy	68	KLOKKNKQRLRNDPLNOK 85	
Db	61	KLOKKNKQRLRNDPLNOK 78	
RESULT 12			
ID	09V424	PRELIMINARY;	PRT; 226 AA.
AC	09V424;		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-MAY-2002 (Tremblrel. 20, Last annotation update)		
DE	Methyl-CpG-binding-domain-like-protein.		
GN	Methyl-CpG-BINDING-DOMAIN-LIKE-PROTEIN OR CG8208.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
NCBI_TaxId=7227;			
UN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN-BERKELEY.		
RC	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Bridton R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Man K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,		
RA	Abbil J.F., Agbayani A., An H.-J., Andrews-Piankocch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,		
RA	De Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Gang N.S., Gelbart W.M., Glasser K.,		
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,		
RA	Jajani M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kraevitz S., Kulp D., Lai Z.,		
RA	Lammel P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wasserman D.A., Westlock G.M., Weisenbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu Q.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster.";		
RL	Science 287:2185-2195(2000).		
RM	[12]		
RP	SEQUENCE FROM N.A.		
RA	Wade P.A., Geggione A., Jones P.L., Ballestar E., Aubry F.,		
RA	Wolffe A.P.;		
RT	"The M-2 histone deacetylase complex couples DNA methylation to		
RT	chromatin remodeling and histone deacetylation.";		
RL	Nat. Genet. 0:0-0(1999).		
DR	EMBL: AE003683; AA544401.1; -		
DR	EMBL: AF111098; AA055391.1; -		
DR	F1Y45e; FBgn0027950; methyl-CpG-binding-domain-like-protein.		

Query Match	28.6%	Score 384.5;	DB 5;	Length 226;
Best Local Similarity	38.3%	Pred. No. 2,1e-23;		
Matches 93;	Conservative 43;	Mismatches 70;	Indels 37;	Gaps 7;
QY	1	MDCALPPGKMKKEVIRKSLGAGSDVYYPSSGKKRFRKPOLARYLGNITVLSSEDFR	60	
DB	14	VDCSVLPFGWQORDEY-KRSGSSANN-----NASNNNSASTASSNNNNNNVVDFY---	63	
QY	61	TGKMPKSLQKKNKORLNDPLNQNKGKPDLTTLPIROTASIFKOPVTYKTNHPSN--KV	118	
DB	64	-----SRALRT-----DVLVLPPIROTASIFKOPVTYKTNHPSN-----	101	
QY	119	KSDPQR-WNEQPROLFWEKRLQGLASDVTEQIITKTMEJLPRGLQGVGSGNDETLISAVA	177	
DB	102	KNEPHGTRERPKQFLWEKRLERLRACHDSEELDDISLPTIRKGVNVEQYVLSVA	161	
QY	178	SALHTSSAPITGVGS--AAVEKNPAWMLNTSOPLCARFVTEDEDIRKOEERYQVQRKLE	235	
DB	162	TALHMLNNGVHGQGSTKDLTKNMAFAMNPQPLMHAVIISSEDIRKQEDRVGARRRKIQ	221	
QY	236	EAL 238		
DB	222	DAL 224		
RESULT 13				
Q8MWY6				
ID	Q8MWY6	PRELIMINARY;	PRT;	194 AA.
AC	Q8MWY6			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Methyl-Cpg binding domain protein 3-like.			
GN	MBD3L.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI.TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Jiang C.L., Szabo P.E., O'Connor T.R., Mann J.R., Pfeiffer G.P.;			
RT	"MBD3L, a gene homologous to the methyl-Cpg-binding domain protein			
RT	genes MBD3 and MBD2, is expressed specifically in postmeiotic cells of			
RL	the testis."			
DR	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY038022; AAK72591.1; "			
SO	SEQUENCE 194 AA; 21562 MW; F9EC132569145DC9 CRC64;			
Query Match	25.2%	Score 339;	DB 4;	Length 194;
Best Local Similarity	41.8%	Pred. No. 8.4e-20;		
Matches 76;	Conservative 39;	Mismatches 65;	Indels 2;	Gaps 2;
QY	69	LQNKQRLRNDPLNQNKGKPDLTTLPIROTASIFKOPVTYKTNHPSNKKVSKSDP-QRNE	127	
DB	1	MAKSSQRRQRODCVNCCKKPGSLSTIPLMSYTRKRPRTTRTPRGNVEVRYHMEESLE	60	
QY	128	OPROLFWERKLGSLASDVTEQIITKTMEJLPRGLQGVGSGNDETLISAVASAL-HTSNP	186	
DB	61	KFOQCMQRRRLQGLQAVSSAGELSLTDLANTLQKLVPSYTGSSLLLEDLSGLEHSCMP	120	
QY	167	ITGVSAVEKNPAWMLNTSOPLCARFVTEDEDIRKOEERYQVQRKLEALMADILSRA	246	
DB	121	HLACSSDAVEIIPAGVGISQLCKQFLVTEEDIRKQESKVTYVERLAIALIADGLANE	180	
QY	247	AD 248		
DB	181	AE 182		

```

RESULT 14
O9D253 ID O9D253 PRELIMINARY: PRT: 186 AA.
AC O9D253
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 1700095H13Rik protein.
GN 1700095H13Rik
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Araiwa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuk S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK018931; BAB31487.1;
DR MGD: MGI:1925722; 1700095H13Rik.
SQ SEQUENCE 186 AA; 20659 MW; EF450ABF69D2AF0 CRC64;

Query Match 18.3%; Score 246; DB 11; Length 186;
Best Local Similarity 35.9%; Pred. No. 2.7e-12;
Matches 66; Conservative 36; Mismatches 68; Indels 14; Gaps 5;

OY 71 KKKORLNDPLNONGKRPDLNTTLPIROTASIFKOPVTKYVTHNPSNKKV-SDPQRNDEP 129
DB 3 KTSQKQCD--CENPSKPCLSITSLPMSYTFKRVRVTKITSHLGNEVRYQWETLEKP 60
OY 130 ROLFWEKRLGSLASDVTEIITKTMELPKLOGVPGSNDLTLASVASALHT--SSAPI 187
DB 61 EQASQKRLGLOAVSSAGELLSTSLAKTLK-----DLSTDPYASASDPTQNTSIDI 113
OY 188 TGOVSAVEKNPAVWLNTSOP--LCKAFIVTDEDIRKOEERVOQVRKKLEBALMADILSR 245
DB 114 TSVPTLESSSHLANMIPKAGPOLCKEFLVTEODIINQERVKYIARERLAVALLAHKLAS 173
OY 246 AADT 249
DB 174 EMET 177

```

```

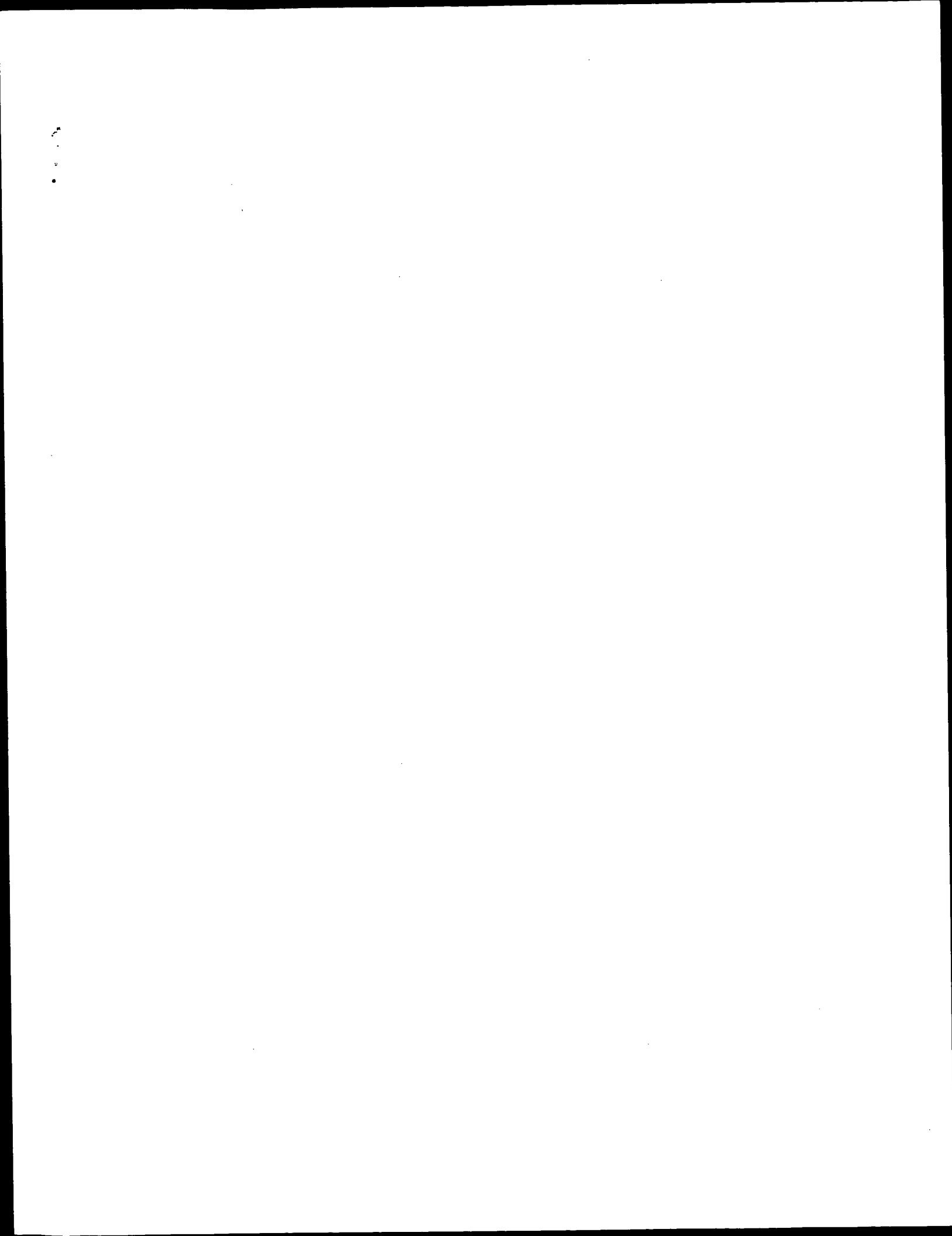
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Araiwa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuk S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Jiang C.L., Szabo P.E., O'Connor T.R., Mann J.R., Pfeiffer G.P.;
RT "MBD3L, a gene homologous to the methyl-CpG-binding domain protein
RT genes MBD3 and MBD2, is expressed specifically in postmeiotic cells of
RT the testis.";
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK06927; BAB24794.1;
DR MBL: AY038023; AAK72592.1;
DR MGD: MGI:1920753; 1700070G05Rik.
SQ SEQUENCE 186 AA; 20675 MW; F5F45BABB69D2AF0 CRC64;

Query Match 18.3%; Score 246; DB 11; Length 186;
Best Local Similarity 35.9%; Pred. No. 2.7e-12;
Matches 66; Conservative 36; Mismatches 68; Indels 14; Gaps 5;

OY 71 KKKORLNDPLNONGKRPDLNTTLPIROTASIFKOPVTKYVTHNPSNKKV-SDPQRNDEP 129
DB 3 KTSQKQCD--CENPSKPCLSITSLPMSYTFKRVRVTKITSHLGNEVRYQWETLEKP 60
OY 130 ROLFWEKRLGSLASDVTEIITKTMELPKLOGVPGSNDLTLASVASALHT--SSAPI 187
DB 61 EQASQKRLGLOAVSSAGELLSTSLAKTLK-----DLSTDPYASASDPTQNTSIDI 113
OY 188 TGOVSAVEKNPAVWLNTSOP--LCKAFIVTDEDIRKOEERVOQVRKKLEBALMADILSR 245
DB 114 TSVPTLESSSHLANMIPKAGPOLCKEFLVTEODIINQERVKYIARERLAVALLAHKLAS 173
OY 246 AADT 249
DB 174 EMET 177

```

Search completed: March 12, 2003, 09:12:43
 Job time : 39.2051 secs



GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2003, 03:17:33 ; Search time 56.7949 Seconds
(without alignments)
1491.072 Million cell updates/sec

Title: US-09-554-414b-2

Perfect score: 2167
Sequence: 1 MRAHPGCGRCPEDEGEESA.....LSRAADTEEMDIEMDSDEA 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp-unclassified:*
- 15: sp-virus:*
- 16: sp-bacteriap:*
- 17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2167	100.0	411	4 Q9UBB5	Q9UBB5 homo sapien
2	2091.5	96.5	414	11 Q9Z2E1	Q9Z2E1 mus musculu
3	1274	58.8	302	4 Q9UIS8	Q9UIS8 homo sapien
4	1210.5	55.9	249	11 Q9Z2D9	Q9Z2D9 mus musculu
5	1059	48.9	282	13 Q9PUM8	Q9PUM8 xenopus lae
6	1056.5	48.8	303	13 Q9PUM8	Q9PUM8 xenopus lae
7	1030	47.5	291	4 Q9S983	Q9S983 homo sapien
8	1021.5	47.1	285	11 Q9Z2D8	Q9Z2D8 mus musculu
9	1013	46.7	200	4 Q60535	Q60535 homo sapien
10	452.5	20.9	314	5 Q9VH87	Q9VH87 drosophila
11	409	18.9	146	4 Q9S242	Q9S242 homo sapien
12	400.5	18.5	226	5 Q9Y424	Q9Y424 drosophila
13	339	15.6	194	4 Q8WY6	Q8WY6 homo sapien
14	271	12.5	619	12 Q9IP09	Q9IP09 cynomolgus
15	264.5	12.2	588	12 Q9IP08	Q9IP08 cynomolgus
16	246	11.4	186	11 Q9D283	Q9D283 mus musculu

17	246	11.4	186	11 Q9D9H3	Q9D9H3 mus musculu
18	243.5	11.2	511	12 Q91332	Q91332 cercopithec
19	237.5	11.0	588	11 Q9DC19	Q9DC19 mus musculu
20	230.5	10.6	636	11 Q9Z2E2	Q9Z2E2 mus musculu
21	226	10.4	476	12 Q80890	Q80890 herpesvirus
22	225	10.4	592	16 Q9PF60	Q9PF60 xyella fas
23	220	10.2	221	10 Q65514	Q65514 arabidopsis
24	218.5	10.1	610	5 Q9VSV8	Q9VSV8 drosophila
25	217.5	10.0	385	5 Q93424	Q93424 caenorhabd
26	216	10.0	175	10 Q9LSN6	Q9LSN6 arabidopsis
27	214.5	9.9	277	4 Q96EC1	Q96EC1 homo sapien
28	214.5	9.9	503	4 Q9UNZ6	Q9UNZ6 homo sapien
29	214.5	9.9	549	4 Q9UNZ7	Q9UNZ7 homo sapien
30	214.5	9.9	556	4 Q15248	Q15248 homo sapien
31	214.5	9.9	586	4 Q9UNZ8	Q9UNZ8 homo sapien
32	214.5	9.9	605	4 Q9UNZ9	Q9UNZ9 homo sapien
33	214.5	9.9	605	4 Q9UIS9	Q9UIS9 homo sapien
34	214	9.9	271	10 Q08529	Q08529 nicotiana t
35	211.5	9.8	291	10 Q39337	Q39337 brassica na
36	209	9.6	344	13 Q42403	Q42403 gallus gall
37	208.5	9.6	396	10 Q65450	Q65450 arabidopsis
38	207.5	9.6	174	10 Q9LTP5	Q9LTP5 arabidopsis
39	207	9.6	255	10 Q9S1H2	Q9S1H2 arabidopsis
40	205.5	9.5	486	6 Q951G8	Q951G8 macaca fasc
41	203	9.4	1432	10 Q9FPR8	Q9FPR8 chlamydomon
42	202.5	9.3	302	10 Q9SL09	Q9SL09 arabidopsis
43	202	9.3	293	13 Q9DEX9	Q9DEX9 cyprinus ca
44	201	9.3	147	5 Q24348	Q24348 drosophila
45	200.5	9.3	173	10 Q41191	Q41191 arabidopsis

ALIGNMENTS

RESULT 1

Q9UBB5 ID Q9UBB5 PRELIMINARY; PRT; 411 AA.

AC Q9UBB5; 01-MAY-2000 (TREMURel. 13, Created)

DT 01-MAY-2000 (TREMURel. 13, Last sequence update)

DT 01-MAY-2002 (TREMURel. 20, Last annotation update)

DE Methyl-CpG binding protein 2.

OS MBD2.

GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99373255; PubMed=10441743;

RA Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;

RT "Genomic structure and chromosomal mapping of the murine and human

RT mbd1, mbd2, mbd3, and mbd4 genes.";

RL Mamm. Genome 10:906-912(1999).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=98449942; PubMed=9774669;

RA Hendrich B., Bird A.;

RT "Identification and characterization of a family of mammalian methyl-

RT CpG binding proteins.";

RL Mol. Cell. Biol. 18:6538-6547(1998).

DR EMBL: AF120983; AAD56597.1; -

DR EMBL: AF120989; AAD56597.1; JOINED.

DR EMBL: AF120990; AAD56597.1; JOINED.

DR EMBL: AF120991; AAD56597.1; JOINED.

DR EMBL: AF120992; AAD56597.1; JOINED.

DR EMBL: AF072242; AAC68871.1; -

DR InterPro: IPR001739; Methyl-CpG_bind.

DR Pfam: PF01428; MBD. 1.

DR SMART: SM00391; MBD. 1.

SO SEQUENCE 411 AA; 43254 MW; FC45E0C9BA0FFA CRC64;

Query Match 100.0%; Score 2167; DB 4; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1.2e-152;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAHGGGRCCEOEESGAAGSGAGDSALIEGCGSALAPSPVSGVRRREGARGRG 60
 |||||
 Db 1 MRAHPGGGRCCEOEESGAAGSGAGDSALIEGCGSALAPSPVSGVRRREGARGRG 60

QY 61 RGRMKOAGRGGVCGR 120
 |||||
 Db 61 RGRMKOAGRGGVCGR 120

QY 121 APRREVPFPPSGSAGCPGPRGRATESGKRMDCPALPQMKKEVIRKSGLSAGKSDVYF 180
 |||||
 Db 121 APRREVPFPPSGSAGCPGPRGRATESGKRMDCPALPQMKKEVIRKSGLSAGKSDVYF 180

QY 181 SPFGKFKFRSKPOLARYLGNTVLDSSFDFTGKMPKSKLQKNKORLRNDPLNOKGKPD 240
 |||||
 Db 181 SPFGKFKFRSKPOLARYLGNTVLDSSFDFTGKMPKSKLQKNKORLRNDPLNOKGKPD 240

QY 241 TLPLIRQTASIFKQPVTKVTHNPSNKVKSQDQRMNEQPRQLFWEKRLQGLSASDVTE 300
 |||||
 Db 241 TLPLIRQTASIFKQPVTKVTHNPSNKVKSQDQRMNEQPRQLFWEKRLQGLSASDVTE 300

QY 301 KTMELPKGLQGVGSGNDETLISAVASALHTSSAPITGOVSAAYEKNPAVWLNTSOP 360
 |||||
 Db 301 KTMELPKGLQGVGSGNDETLISAVASALHTSSAPITGOVSAAYEKNPAVWLNTSOP 360

QY 361 AFIVTDEDIRKQEEVVOVRKKLEALMADILSRAADTEMDIEMDSGDEA 411
 |||||
 Db 361 AFIVTDEDIRKQEEVVOVRKKLEALMADILSRAADTEMDIEMDSGDEA 411

RESULT 2

Q922E1 PRELIMINARY; PRT; 414 AA.

AC Q922E1;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Methyl-Cpg binding protein MBD2.
 GN MBD2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=98449942; PubMed=9774669;
 RA Hendrich B., Bird A.;
 RT "Identification and characterization of a family of mammalian methyl-
 CpG binding proteins."
 RL Moll. Cell. Biol. 18:6538-6547(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=99373255; PubMed=10441743;
 RA Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;
 RT "Genomic structure and chromosomal mapping of the murine and human
 Mbd1, mbd2, mbd3, and mbd4 genes."
 RL Mamm. Genome 10:906-912(1999).
 DR EMBL: AF072243; AAC68872.1; -;
 DR EMBL: AF120986; AAD50372.1; -;
 DR EMBL: AF120983; AAD50372.1; JOINED.
 DR EMBL: AF120984; AAD50372.1; JOINED.
 DR EMBL: AF120985; AAD50372.1; JOINED.
 DR MGD; MGI:133813; Mbd2.
 DR InterPro: IPR001739; Methyl-Cpg_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SQ SEQUENCE 414 AA; 43543 MW; 9601D95E347E8E53 CRC64;

Query Match 96.5%; Score 2091.5; DB 11; Length 414;
 Best Local Similarity 96.6%; Pred. No. 4.7e-147;
 Matches 400; Conservative 4; Mismatches 7; Indels 3; Gaps 1;

QY 1 MRAHGGGRCCEOEESGAAGSGAGDSALIEGCGSALAPSPVSGVRRREGARGRG 60
 |||||
 Db 1 MRAHPGGGRCCEOEESGAAGSGAGDSALIEGCGSALAPSPVSGVRRREGARGRG 60

QY 61 RGRMKOAGRGGVCGR 117
 |||||
 Db 61 RGRMKOAGRGGVCGR 120

QY 118 GCGAPREVPFPPSGSAGCPGPRGRATESGKRMDCPALPQMKKEVIRKSGLSAGKSDV 177
 |||||
 Db 121 GCGAPREVPFPPSGSAGCPGPRGRATESGKRMDCPALPQMKKEVIRKSGLSAGKSDV 180

QY 178 YFSPSGKFKFRSKPOLARYLGNTVLDSSFDFTGKMPKSKLQKNKORLRNDPLNOKGK 237
 |||||
 Db 181 YFSPSGKFKFRSKPOLARYLGNTVLDSSFDFTGKMPKSKLQKNKORLRNDPLNOKGK 240

QY 238 DLNTTLIRQTASIFKQPVTKVTHNPSNKVKSQDQRMNEQPRQLFWEKRLQGLSASDV 297
 |||||
 Db 241 DLNTTLIRQTASIFKQPVTKVTHNPSNKVKSQDQRMNEQPRQLFWEKRLQGLSASDV 300

QY 298 QITMELPKGLQGVGSGNDETLISAVASALHTSSAPITGOVSAAYEKNPAVWLNTSOP 357
 |||||
 Db 301 QITMELPKGLQGVGSGNDETLISAVASALHTSSAPITGOVSAAYEKNPAVWLNTSOP 360

QY 358 LCKAFIVTDEDIRKQEEVVOVRKKLEALMADILSRAADTEMDIEMDSGDEA 411
 |||||
 Db 361 LCKAFIVTDEDIRKQEEVVOVRKKLEALMADILSRAADTEMDIEMDSGDEA 414

RESULT 3

Q90IS8 PRELIMINARY; PRT; 302 AA.

AC Q90IS8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Testis-specific methyl-Cpg binding protein 2.
 GN MBD2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99373255; PubMed=10441743;
 RA Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;
 RT "Genomic structure and chromosomal mapping of the murine and human
 Mbd1, Mbd2, Mbd3, and Mbd4 genes."
 RL Mamm. Genome 10:906-912(1999).
 DR EMBL: AF120989; AAD56596.1; -;
 DR EMBL: AF120988; AAD56596.1; JOINED.
 DR InterPro: IPR001739; Methyl-Cpg_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SQ SEQUENCE 302 AA; 31745 MW; CCEC65D926222717 CRC64;

Query Match 58.8%; Score 1274; DB 4; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.3e-86;
 Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAHGGGRCCEOEESGAAGSGAGDSALIEGCGSALAPSPVSGVRRREGARGRG 60
 |||||
 Db 1 MRAHPGGGRCCEOEESGAAGSGAGDSALIEGCGSALAPSPVSGVRRREGARGRG 60

QY 61 RGRMKOAGRGGVCGR 120
 |||||
 Db 61 RGRMKOAGRGGVCGR 120

QY 121 APRREVPFPPSGSAGCPGPRGRATESGKRMDCPALPQMKKEVIRKSGLSAGKSDVYF 180

Db 121 APRREPVPFSGSAGPGRPRATESGKRMDCALPGRMKKEVIRKSGLSAKSDVYTF 180
 181 SPSSKFRKSRKPOLARYLGNVDSSEPFRRGKMKPSKLOKKNKRLNDPLNOK 234
 181 SPSSKFRKSRKPOLARYLGNVDSSEPFRRGKMKPSKLOKKNKRLNDPLNOK 234

RESULT 4

092ZD9 PRELIMINARY; PRT; 249 AA.
 AC 092ZD9;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Testis specific methyl-CpG binding protein MBD2.
 GN MBD2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 1)
 RC SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6; TISSUE=TESTIS;
 RX MEDLINE=98449942; PubMed=9774669;
 RT Hendrich B., Bird A.;
 RT "Identification and characterization of a family of mammalian methyl-CpG binding proteins.";
 RT Mol. Cell. Biol. 18:6538-6547(1998).
 RL 12)
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=99373255; PubMed=10441743;
 RX Hendrich B., Abbott C., McQueen H., Chambers D., Cross S., Bird A.;
 RT "Genomic structure and chromosomal mapping of the murine and human mbd1, mbd2, mbd3, and mbd4 genes.";
 RT Mamm. Genome 10:906-912(1999).
 RL EMBL: AF072245; AAC68874.1; -;
 DR EMBL: AF120983; AAD50373.1; -;
 DR MGI: 133813; Mbd2.
 DR InterPro: IPR001739; Methyl-CpG_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SQ SEQUENCE 249 AA; 25493 MW; 530D751EBE5C9540 CRC64;

Query Match 55.9%; Score 1210.5; DB 11; Length 249;
 Best Local Similarity 95.4%; Pred. No. 5e-82;

Matches 226; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 1 MRAHREGGKCCPEOEGESAGSGAGSDAIEGGGGSALAPSPVSGVRRREGARGGGRG 60
 DB 1 MRAHREGGKCCPEOEGESAGSGAGSDAIEGGGGSALAPSPVSGVRRREGARGGGRG 60
 QY 61 RGRMKOAGRGVCGVRGR 117
 DB 61 RGRMKOAGRGVCGVRGR 117
 QY 118 GGGADPREVPSPSGSAGPGRPRATESGKRMDCPALPGRMKKEVIRKSGLSAKSDV 177
 DB 118 GGGADPREVPSPSGSAGPGRPRATESGKRMDCPALPGRMKKEVIRKSGLSAKSDV 177
 QY 121 GGVARRDPVPFPPSSSSSGPGRPRATESGKRMDCPALPGRMKKEVIRKSGLSAKSDV 180
 DB 121 GGVARRDPVPFPPSSSSSGPGRPRATESGKRMDCPALPGRMKKEVIRKSGLSAKSDV 180
 QY 178 YFSSSGKKFRKSRKPOLARYLGNVDSSEPFRRGKMKPSKLOKKNKRLNDPLNOK 234
 DB 178 YFSSSGKKFRKSRKPOLARYLGNVDSSEPFRRGKMKPSKLOKKNKRLNDPLNOK 234
 QY 181 YFSSSGKKFRKSRKPOLARYLGNVDSSEPFRRGKMKPSKLOKKNKRLNDPLNOK 237
 DB 181 YFSSSGKKFRKSRKPOLARYLGNVDSSEPFRRGKMKPSKLOKKNKRLNDPLNOK 237

RESULT 5

09PUM9 PRELIMINARY; PRT; 282 AA.
 AC 09PUM9;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Methyl-CpG binding protein MBD3.
 GN MBD3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN 1)
 RP SEQUENCE FROM N.A.
 RA Wade P.A., Gegonne A., Jones P.L., Ballestar E., Aubry F.,
 RA Wolffe A.P.;
 RT "The M1-2 histone deacetylase complex couples DNA methylation to
 RT chromatin remodeling and histone deacetylation.";
 RL Nat. Genet. 0:0-0(1999).
 DR EMBL: AF170347; AAD55390.1; -;
 DR InterPro: IPR001739; Methyl-CpG_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SQ SEQUENCE 303 AA; 34078 MW; 6891EFA05B0371E9 CRC64;

Query Match 48.9%; Score 1059; DB 13; Length 282;
 Best Local Similarity 75.4%; Pred. No. 1e-70;
 Matches 202; Conservative 33; Mismatches 27; Indels 6; Gaps 2;

QY 148 KRMDPALPGRMKKEVIRKSGLSAKSDVYFSPSGKFRKSRKPOLARYLGNVDSSE 207
 DB 4 KRMECSAL-QGKKKEVIRKSGLSAKSDVYFSPSGKFRKSRKPOLARYLGNVDSSE 207
 QY 208 FRFGKMKPSKLOKKNKRLNDPLNOKKGRPDNLNTPLPRTGASIFKQPVTKVTHNPSKV 267
 DB 63 FRFGKMKPSKLOKKNKRLNDPLNOKKGRPDNLNTPLPRTGASIFKQPVTKVTHNPSKV 267
 QY 268 KSDPQRMNEQROQLFEKRLGSLASDVTEQITKTMELPKGLGCVGSGNDETLASVAS 327
 DB 123 KSDPQRMNEQROQLFEKRLGSLASDVTEQITKTMELPKGLGCVGSGNDETLASVAS 327
 QY 328 ALHTSSAPITGQVSAVAEKNPAVWLNTSOPCKRAFIYTDIEDIRKQEEBRYOVRKLEAL 387
 DB 183 ALHTSSAPITGQVSAVAEKNPAVWLNTSOPCKRAFIYTDIEDIRKQEEBRYOVRKLEAL 387
 QY 388 MADLISRAADTEE-----MDIEMDSGDE 410
 DB 243 MADLHAHEEISKDGAPLDKIDDEE 270

RESULT 6

09PUM8 PRELIMINARY; PRT; 303 AA.
 AC 09PUM8;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Methyl-CpG binding protein MBD3 long form.
 GN MBD3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN 1)
 RP SEQUENCE FROM N.A.
 RA Wade P.A., Gegonne A., Jones P.L., Ballestar E., Aubry F.,
 RA Wolffe A.P.;
 RT "The M1-2 histone deacetylase complex couples DNA methylation to
 RT chromatin remodeling and histone deacetylation.";
 RL Nat. Genet. 0:0-0(1999).
 DR EMBL: AF170347; AAD55390.1; -;
 DR InterPro: IPR001739; Methyl-CpG_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00391; MBD; 1.
 SQ SEQUENCE 303 AA; 34078 MW; 6891EFA05B0371E9 CRC64;

Query Match 48.8%; Score 1056.5; DB 13; Length 303;


```

Oy 388 MADLISRAADTEEM 401
| | | | |
Db 244 MADML---AHVEL 254

RESULT 9
O60535 PRELIMINARY: PRT: 200 AA.
ID 060535
AC 060535
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
GN Antigen NY-CO-41 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON CANCER METASTASIS TO LUNG;
RX MEDLINE=98272252; PubMed=9610721;
RA Scanlan M.J., Chen Y.T., Williamson B., Gure A.O., Stockert E.,
RA Gordon J.D., Tureci O., Sahin U., Pfreundschuh M., Old L.J.;
RT "Characterization of human colon cancer antigens recognized by
RT autologous antibodies."
RL Int. J. Cancer 76:652-658(1998).
DR EMBL: AF039701; AAC18050.1; -.
FT NON TER
SO SEQUENCE 200 AA; 22270 MW; A70B46B35CA55A5D CRC64;

Query Match 46.7%; Score 1013; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 1,6e-67;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 212 KMSPSKLOKKNORLNDPLNKNKRPDLNTLPIRQASIFKQDVTYVTHNPKVKSDP 271
| | | | |
Db 1 KMSPSKLOKKNORLNDPLNKNKRPDLNTLPIRQASIFKQDVTYVTHNPKVKSDP 60

Oy 272 QRMNEOPROLFEWKRLOGLSASDVTEIOIKTMELPKGLOGGSPSNDETLISAVASLHT 331
| | | | |
Db 61 QRMNEOPROLFEWKRLOGLSASDVTEIOIKTMELPKGLOGGSPSNDETLISAVASLHT 120

Oy 332 SSATITGOVSAVKNPAVWINTSOPLCARFIVTDEDIRKQERVOQVRKKLEALMADI 391
| | | | |
Db 121 SSATITGOVSAVKNPAVWINTSOPLCARFIVTDEDIRKQERVOQVRKKLEALMADI 180

Oy 392 LSRAADTEEMDIEMDSGEA 411
| | | | |
Db 181 LSRAADTEEMDIEMDSGEA 200

RESULT 10
O9VHB7 PRELIMINARY: PRT: 314 AA.
ID 09VHB7
AC 09VHB7
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE Mch1-CPG-binding-domain-like-protein.
GN MCH1L-CPG-BINDING-DOMAIN-LIKE-PROTEIN OR CG6208.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceinalker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthakrishnan P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

```

```

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miles G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahler C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glisler C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Klamm B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
SC Science 287:2185-2195(2000).
DR EMBL: AE003683; AAF54400.1; -.
DR Flybase: FBgn0027950; methyl-CpG-binding-domain-like-protein.
DR Interpro: IPR001739; Methyl-CpG_bind.
DR Pfam: PF01429; MBD; 1.
DR SMART: SM00391; MBD; 1.
SO SEQUENCE 314 AA; 33803 MW; 7B6BDFE63873D230 CRC64;

Query Match 20.9%; Score 452.5; DB 5; Length 314;
Best Local Similarity 35.2%; Pred. No. 1e-25;
Matches 112; Conservative 48; Mismatches 77; Indels 81; Gaps 9;

Oy 141 PRATESGRMDCPALPPGKKKEVIRKSGLSA-----GKSDVYVF 180
| | | | |
Db 5 PSVTIERKRVDCSVLPKGMQREY--RKSSGSSANNNASSNNSSATASSNNNNKKYDVFY 63

Oy 181 SPGKKFRSKPOLARLGTIVLSSDFRTGK-----MP----- 215
| | | | |
Db 64 SPGKKRAEGKPO-----DIAIPDFQPKMPHCALPSPISLYCSAMPLPIASGG 114

Oy 216 -----SKLOKKORLNDPLNONGK-----KPDNTLPIRQTSIRKO 254
| | | | |
Db 115 NGATSSANNAALKRKFARSGGNAAGAAPPAATASBALFTDVLVPIRQTASIRKO 174

Oy 255 PVTYVTHNPSN--KVSDFQR--MNEOPROLFEWKRLOGLSASDVTEIOIKTMELPKGLOG 311
| | | | |
Db 175 PVTYVIRNHKODPAKAKREKHGTREKPKOLEWKRLERLRACHSGEGLDLSLPKTI RT 234

Oy 312 VPGSNDERTLSAVASLHTSAPITGOVS--AAVEKPAVWINTSOPLCARFIVTDEDI 369
| | | | |
Db 235 VGPVNVNEQIVLOSATVATLHMLNAGVHSGSSTKADLTKNAMAFMNEOPLMHAIVISED 294

Oy 370 RKQERVOQVRKKLEAL 387
| | | | |
Db 295 RKQEDRVGVARRKQLDAL 312

RESULT 11
O95242

```

```

ID 095242 PRELIMINARY; PRT; 146 AA.
AC 095242;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Testis specific methyl-CpG binding protein MBD2 (Fragment).
GN MBD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=98449942; PubMed=9774669;
RA Hendlich B., Bird A.;
RT "Identification and characterization of a family of mammalian methyl-
RT CpG binding proteins."
RL Mol. Cell. Biol. 18:6538-6547(1998).
DR EMBL; AF072246; AAC68673.1; -
DR InterPro; IPR001739; Methyl-CpG_bind.
DR Pfam; PF01429; MBD; 1.
DR SMART; SM00391; MBD; 1.
FT NON_TER
SQ SEQUENCE 146 AA; 16931 MW; 9D6CC3CFE140EEEA CRC64;

Query Match 18.9%; Score 409; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 6.5e-23;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 PGMKKEEVIRKSGLSAGKSVYFSPGKFRSKPOLARYLGMTVDSDFRTGKMPS 216
DB 1 PGMKKEEVIRKSGLSAGKSVYFSPGKFRSKPOLARYLGMTVDSDFRTGKMPS 60
QY 217 KLOKNKQRLRNDPLNQK 234
DB 61 KLOKNKQRLRNDPLNQK 78

RESULT 12
ID 09V424 PRELIMINARY; PRT; 226 AA.
AC 09V424;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Methyl-CpG-binding-domain-like-protein.
GN METHYL-CpG-BINDING-DOMAIN-LIKE-PROTEIN OR CG8208.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephyridia; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731123;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintinas S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

```

```

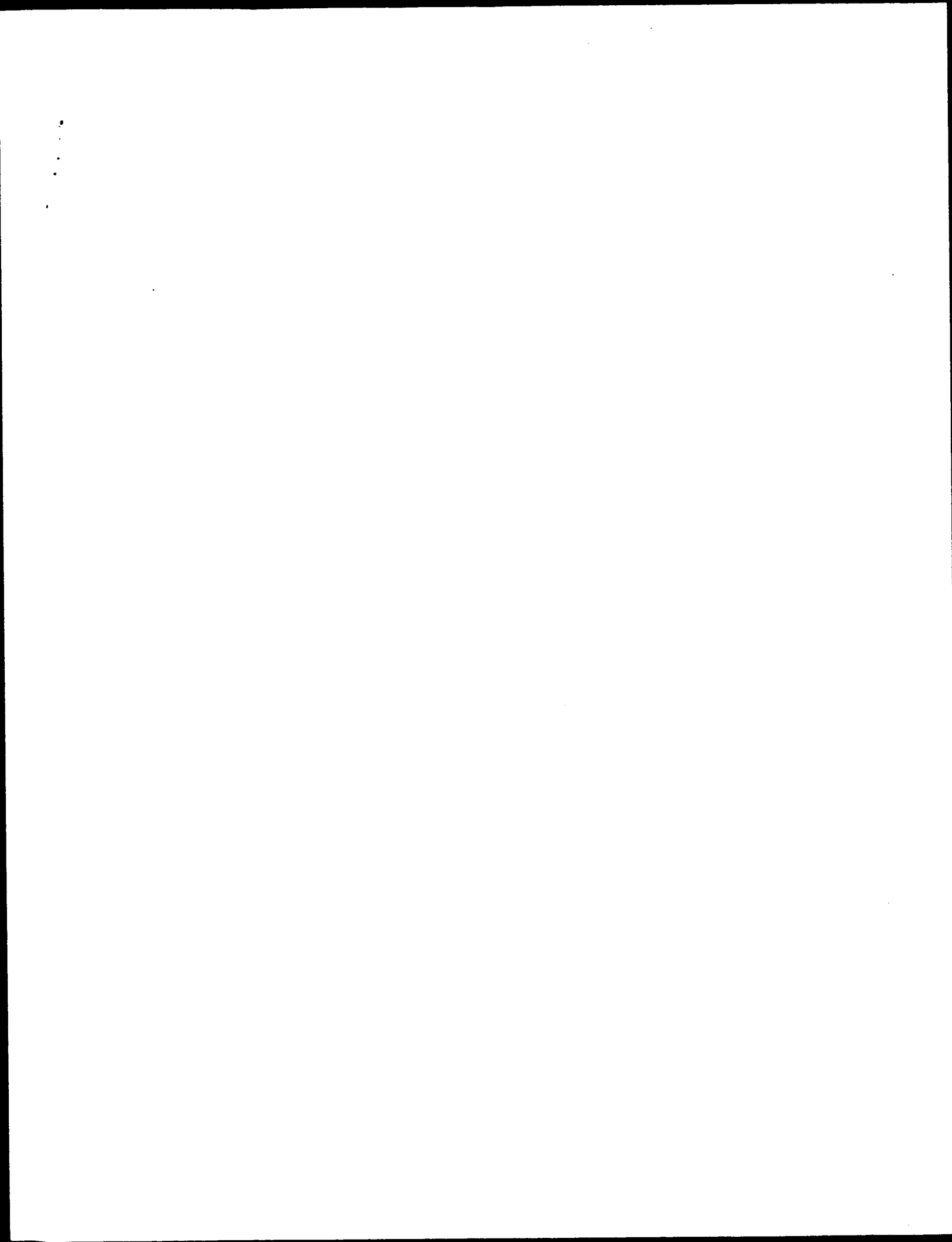
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskens D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN 121
RP SEQUENCE FROM N.A.
RA Wade P.A., Geggione A., Jones P.L., Ballestar E., Aubry F.,
RA Wolffe A.P.;
RT "The M1-2 histone deacetylase complex couples DNA methylation to
RT chromatin remodeling and histone deacetylation."
RL Nat. Genet. 0:0-0(1999).
RL EMBL; AE003683; AAF54401.1; -
DR EMBL; AF171098; AAD53391.1; -
DR Flybase; FBgn027950; methyl-CpG-binding-domain-like-protein.
DR InterPro; IPR001739; Methyl-CpG_bind.
DR Pfam; PF01429; MBD; 1.
DR SMART; SM00391; MBD; 1.
SQ SEQUENCE 226 AA; 25292 MW; 26F2DD296442B4D3 CRC64;

Query Match 18.5%; Score 400.5; DB 5; Length 226;
Best Local Similarity 38.5%; Pred. No. 4.9e-22;
Matches 97; Conservative 43; Mismatches 75; Indels 37; Gaps 7;

QY 141 PRATSGRMDCPALPPGKKEEVIRKSGLSAGKSVYFSPGKFRSKPOLARYLGMT 200
DB 5 PSYTIERRKVDSCVLPKGMQROEV-RKSGSSANN-----NASSNNNSATSSNNNNK 57
QY 201 VDSLSDFRTGKMPSKLOKNKQRLRNDPLNQKGPDLNTLPIRQASIFKQPVYVT 260
DB 58 VDFEYF-----SRALRT-----DVSIVPIRQTFASIFKQPVYIR 92
QY 261 NHPSN--KYKSPQR-MNEQPOLTFEKRLOGLSASDVTEQITKIMELPKGLGVGPSN 317
DB 93 NHPQDPAKKAKEPKHGTREKRPKOLFWEKRLERLRACHDSGEELDDISLTKTIRTVGPVNV 152
QY 318 DEFLTSAVASALHTSSAPITGOVS--AAVEKMPAVLNTSPFLCAFIYTDIDIRKQER 375
DB 153 EGVTLQSVATLALMLNAGVGSSTKADLTAKNAFMFNEQFLMAHVAIISDDIRKQER 212
QY 376 VQVQRRKLEAL 387
DB 213 VGVARRKLDAL 224

RESULT 13
ID 08MWY6 PRELIMINARY; PRT; 194 AA.
AC 08MWY6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Methyl-CpG binding domain protein 3-like.
GN MBD3L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2003, 05:40:46 ; Search time 19.5423 Seconds

(without alignments)
618.801 Million cell updates/sec

Title: US-09-554-414b-2

Perfect score: 2167
Sequence: 1 MRAHGGCGRCCEQEESEA.....LSRAADTEMDIEMSGDEA 411

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCFUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	134.4	62.0	263	4	US-09-149-476-580
2	214.5	9.9	574	4	US-09-079-431B-6
3	214.5	9.9	629	4	US-09-079-431B-4
4	214.5	9.9	630	4	US-09-079-431B-2
5	214	9.9	641	4	US-09-249-585A-3
6	210	9.7	235	2	US-08-529-180B-1
7	209.5	9.7	201	4	US-09-052-995-1
8	209.5	9.7	201	4	US-09-053-003-40
9	209.5	9.7	201	4	US-09-054-281-22
10	209.5	9.7	201	4	US-09-478-948-6
11	196.5	9.1	1136	4	US-08-806-029-9
12	196.5	9.1	1177	1	US-07-609-716-31
13	196.5	9.1	1177	1	US-08-175-155-29
14	196.5	9.1	1177	1	US-08-477-509B-64
15	196.5	9.1	1177	2	US-08-707-237A-35
16	196.5	9.1	1177	2	US-08-482-085B-64
17	196.5	9.1	1177	4	US-08-475-411A-31
18	196.5	9.1	1177	4	US-08-478-029A-71
19	196.5	9.1	1177	4	US-09-444-791A-64
20	195.5	9.0	1332	1	US-07-609-716-41
21	195.5	9.0	1332	4	US-08-475-411A-41
22	195.5	9.0	1332	4	US-08-478-029A-41
23	195	9.0	745	2	US-09-010-928B-28
24	195	9.0	780	2	US-09-010-928B-2
25	193	8.9	784	1	US-07-609-716-48
26	193	8.9	784	4	US-08-475-411A-48
27	193	8.9	784	4	US-08-478-029A-48

28	192.5	8.9	649	1	US-07-609-716-49	Sequence 49, Appl
29	192.5	8.9	649	4	US-08-475-411A-49	Sequence 49, Appl
30	192.5	8.9	649	4	US-08-478-029A-49	Sequence 49, Appl
31	192.5	8.9	718	1	US-08-425-069-2	Sequence 2, Appl1
32	192.5	8.9	718	1	US-08-317-844B-2	Sequence 2, Appl1
33	192.5	8.9	747	3	US-09-034-177-3	Sequence 3, Appl1
34	192.5	8.9	945	1	US-08-089-862-6	Sequence 6, Appl1
35	192.5	8.9	945	1	US-08-587-333-13	Sequence 13, Appl1
36	192.5	8.9	1059	5	PCT-US94-07776-11	Sequence 11, Appl1
37	192.5	8.9	1059	2	US-08-175-155-48	Sequence 48, Appl
38	192.5	8.9	1059	1	US-08-707-237A-54	Sequence 10, Appl
39	192.5	8.9	1059	4	US-08-806-029-10	Sequence 83, Appl
40	192.5	8.9	1101	1	US-08-477-509B-83	Sequence 83, Appl
41	192.5	8.9	1101	3	US-08-482-085B-83	Sequence 19, Appl
42	192.5	8.9	1101	4	US-08-482-085B-83	Sequence 1, Appl1
43	189.5	8.7	651	4	US-09-444-791A-83	Sequence 1, Appl1
44	189.5	8.7	651	4	US-08-556-978B-19	Sequence 2, Appl1
45	189.5	8.7	651	4	US-09-247-806-1	
	189	8.7	1958	1	US-07-945-283-2	

ALIGNMENTS

RESULT 1

US-09-149-476-580

Sequence 580, Application US/09149476

Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P1

CURRENT APPLICATION NUMBER: US/09/149,476

EARLIER FILING DATE: 1998-09-08

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,583

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,618

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,503

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,592

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,581

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,584

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 1997-05-23	500
EARLIER FILING DATE: 1997-05-23	587
EARLIER APPLICATION NUMBER: 1997-05-23	587
EARLIER FILING DATE: 1997-05-23	492
EARLIER APPLICATION NUMBER: 1997-05-23	598
EARLIER FILING DATE: 1997-05-23	613
EARLIER APPLICATION NUMBER: 1997-05-23	562
EARLIER FILING DATE: 1997-05-23	596
EARLIER APPLICATION NUMBER: 1997-05-23	612
EARLIER FILING DATE: 1997-05-23	632
EARLIER APPLICATION NUMBER: 1997-05-23	601
EARLIER FILING DATE: 1997-05-23	580
EARLIER APPLICATION NUMBER: 1997-04-11	568
EARLIER FILING DATE: 1997-04-11	314
EARLIER APPLICATION NUMBER: 1997-04-11	569
EARLIER FILING DATE: 1997-04-11	311
EARLIER APPLICATION NUMBER: 1997-04-11	671
EARLIER FILING DATE: 1997-04-11	674
EARLIER APPLICATION NUMBER: 1997-04-11	669
EARLIER FILING DATE: 1997-04-11	312
EARLIER APPLICATION NUMBER: 1997-04-11	313
EARLIER FILING DATE: 1997-04-11	672
EARLIER APPLICATION NUMBER: 1997-04-11	315
EARLIER FILING DATE: 1997-04-11	974
EARLIER APPLICATION NUMBER: 1997-06-06	886
EARLIER FILING DATE: 1997-08-22	877
EARLIER APPLICATION NUMBER: 1997-08-22	889
EARLIER FILING DATE: 1997-08-22	893
EARLIER APPLICATION NUMBER: 1997-08-22	630
EARLIER FILING DATE: 1997-08-22	878
EARLIER APPLICATION NUMBER: 1997-08-22	662
EARLIER FILING DATE: 1997-08-22	872
EARLIER APPLICATION NUMBER: 1997-08-22	882
EARLIER FILING DATE: 1997-08-22	637
EARLIER APPLICATION NUMBER: 1997-08-22	903
EARLIER FILING DATE: 1997-08-22	888
EARLIER APPLICATION NUMBER: 1997-08-22	879
EARLIER FILING DATE: 1997-08-22	880
EARLIER APPLICATION NUMBER: 60/056, 880	

1	EARLIER	FILING DATE:	1997-08-22	894
2	EARLIER	APPLICATION NUMBER:	60/056, 894	
3	EARLIER	FILING DATE:	1997-08-22	
4	EARLIER	APPLICATION NUMBER:	60/056, 911	
5	EARLIER	FILING DATE:	1997-08-22	
6	EARLIER	APPLICATION NUMBER:	60/056, 636	
7	EARLIER	FILING DATE:	1997-08-22	
8	EARLIER	APPLICATION NUMBER:	60/056, 874	
9	EARLIER	FILING DATE:	1997-08-22	
10	EARLIER	APPLICATION NUMBER:	60/056, 910	
11	EARLIER	FILING DATE:	1997-08-22	
12	EARLIER	APPLICATION NUMBER:	60/056, 864	
13	EARLIER	FILING DATE:	1997-08-22	
14	EARLIER	APPLICATION NUMBER:	60/056, 631	
15	EARLIER	FILING DATE:	1997-08-22	
16	EARLIER	APPLICATION NUMBER:	60/056, 845	
17	EARLIER	FILING DATE:	1997-08-22	
18	EARLIER	APPLICATION NUMBER:	60/056, 892	
19	EARLIER	FILING DATE:	1997-08-22	
20	EARLIER	APPLICATION NUMBER:	60/057, 761	
21	EARLIER	FILING DATE:	1997-08-22	
22	EARLIER	APPLICATION NUMBER:	60/047, 595	
23	EARLIER	FILING DATE:	1997-05-23	
24	EARLIER	APPLICATION NUMBER:	60/047, 599	
25	EARLIER	FILING DATE:	1997-05-23	
26	EARLIER	APPLICATION NUMBER:	60/047, 598	
27	EARLIER	FILING DATE:	1997-05-23	
28	EARLIER	APPLICATION NUMBER:	60/047, 585	
29	EARLIER	FILING DATE:	1997-05-23	
30	EARLIER	APPLICATION NUMBER:	60/047, 586	
31	EARLIER	FILING DATE:	1997-05-23	
32	EARLIER	APPLICATION NUMBER:	60/047, 590	
33	EARLIER	FILING DATE:	1997-05-23	
34	EARLIER	APPLICATION NUMBER:	60/047, 594	
35	EARLIER	FILING DATE:	1997-05-23	
36	EARLIER	APPLICATION NUMBER:	60/047, 589	
37	EARLIER	FILING DATE:	1997-05-23	
38	EARLIER	APPLICATION NUMBER:	60/047, 593	
39	EARLIER	FILING DATE:	1997-05-23	
40	EARLIER	APPLICATION NUMBER:	60/047, 614	
41	EARLIER	FILING DATE:	1997-05-23	
42	EARLIER	APPLICATION NUMBER:	60/043, 578	
43	EARLIER	FILING DATE:	1997-04-11	
44	EARLIER	APPLICATION NUMBER:	60/043, 576	
45	EARLIER	FILING DATE:	1997-04-11	
46	EARLIER	APPLICATION NUMBER:	60/047, 501	
47	EARLIER	FILING DATE:	1997-05-23	
48	EARLIER	APPLICATION NUMBER:	60/043, 670	
49	EARLIER	FILING DATE:	1997-04-11	
50	EARLIER	APPLICATION NUMBER:	60/056, 632	
51	EARLIER	FILING DATE:	1997-08-22	
52	EARLIER	APPLICATION NUMBER:	60/056, 664	
53	EARLIER	FILING DATE:	1997-08-22	
54	EARLIER	APPLICATION NUMBER:	60/056, 876	
55	EARLIER	FILING DATE:	1997-08-22	
56	EARLIER	APPLICATION NUMBER:	60/056, 881	
57	EARLIER	FILING DATE:	1997-08-22	
58	EARLIER	APPLICATION NUMBER:	60/056, 909	
59	EARLIER	FILING DATE:	1997-08-22	
60	EARLIER	APPLICATION NUMBER:	60/056, 875	
61	EARLIER	FILING DATE:	1997-08-22	
62	EARLIER	APPLICATION NUMBER:	60/056, 862	
63	EARLIER	FILING DATE:	1997-08-22	
64	EARLIER	APPLICATION NUMBER:	60/056, 887	
65	EARLIER	FILING DATE:	1997-08-22	
66	EARLIER	APPLICATION NUMBER:	60/056, 908	
67	EARLIER	FILING DATE:	1997-08-22	
68	EARLIER	APPLICATION NUMBER:	60/048, 964	
69	EARLIER	FILING DATE:	1997-06-06	
70	EARLIER	APPLICATION NUMBER:	60/057, 650	
71	EARLIER	FILING DATE:	1997-09-05	
72	EARLIER	APPLICATION NUMBER:	60/056, 884	
73	EARLIER	FILING DATE:	1997-08-22	

EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 62.0%; Score 1344; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.1e-98;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 MDCPALPGMKKEVIRKSGLSAGKSDVYFSPSGKFRSKPOLARYIGNTVDSLSPDR 209
DB 1 MDCPALPGMKKEVIRKSGLSAGKSDVYFSPSGKFRSKPOLARYIGNTVDSLSPDR 60
QY 210 TGMKPSKLOKNKORLNDPLNOKGKPDNTLPIROTASTIKOPVYKTNHPSKVKVS 269
DB 61 TGMKPSKLOKNKORLNDPLNOKGKPDNTLPIROTASTIKOPVYKTNHPSKVKVS 120
QY 270 DPGRMNOPRQLFMEKRLQGLSADVTETQIKTMEI.PKGLQGVGPGSNDETLISAVASAL 329
DB 121 DPGRMNOPRQLFMEKRLQGLSADVTETQIKTMEI.PKGLQGVGPGSNDETLISAVASAL 180
QY 330 HTSSAPITGVSAVEKNPAWMLNTSOPLCARIVTDEDIRKOEERVOQVRKKLEBALMA 389
DB 181 HTSSAPITGVSAVEKNPAWMLNTSOPLCARIVTDEDIRKOEERVOQVRKKLEBALMA 240
QY 390 DILSRAADTEEMDIEMDSGDEA 411
DB 241 DILSRAADTEEMDIEMDSGDEA 262

RESULT 2

US-09-079-431B-6
Sequence 6, Application US/09079431B
Patent No. 6326484
GENERAL INFORMATION:
APPLICANT: Gage, Fred
APPLICANT: Ueda, Tetsuya
TITLE OF INVENTION: TRANSCRIPTION FACTOR REGULATING FGF-2
FILE REFERENCE: SALKINS.015A
CURRENT APPLICATION NUMBER: US/09/079,431B
CURRENT FILING DATE: 1998-05-14
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 574
TYPE: PRT
ORGANISM: Homo sapiens
US-09-079-431B-6

Query Match 9.9%; Score 214.5; DB 4; Length 574;
Best Local Similarity 45.7%; Pred. No. 4.6e-09;
Matches 42; Conservative 13; Mismatches 30; Indels 7; Gaps 1;

QY 150 MDCPALPGMKKEVIRKSGLSAGKSDVYFSPSGKFRSKPOLARYIGNTVDSLSPDR 209
DB 6 LDCPALPGMKKEVIRKSGATCGRSPTYOSPTGDIRSKVELTRYLGPACDITLDFDK 65
QY 210 TG-----KMPKSKLOKNKORLNDPLNOK 234
DB 66 OGILCYPAKAPHPVAVASKKKRKPAPKTRK 97

RESULT 3

US-09-079-431B-4
Sequence 4, Application US/09079431B
Patent No. 6326484
GENERAL INFORMATION:
APPLICANT: Gage, Fred
APPLICANT: Ueda, Tetsuya
TITLE OF INVENTION: TRANSCRIPTION FACTOR REGULATING FGF-2

TITLE OF INVENTION: AND VARIANTS THEREOF
FILE REFERENCE: SALKINS.015A
CURRENT APPLICATION NUMBER: US/09/079,431B
CURRENT FILING DATE: 1998-05-14
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 629
TYPE: PRT
ORGANISM: Homo sapiens
US-09-079-431B-4

Query Match 9.9%; Score 214.5; DB 4; Length 629;
Best Local Similarity 45.7%; Pred. No. 5.2e-09;
Matches 42; Conservative 13; Mismatches 30; Indels 7; Gaps 1;

QY 150 MDCPALPGMKKEVIRKSGLSAGKSDVYFSPSGKFRSKPOLARYIGNTVDSLSPDR 209
DB 6 LDCPALPGMKKEVIRKSGATCGRSPTYOSPTGDIRSKVELTRYLGPACDITLDFDK 65
QY 210 TG-----KMPKSKLOKNKORLNDPLNOK 234
DB 66 OGILCYPAKAPHPVAVASKKKRKPAPKTRK 97

RESULT 4

US-09-079-431B-2
Sequence 2, Application US/09079431B
Patent No. 6326484
GENERAL INFORMATION:
APPLICANT: Gage, Fred
APPLICANT: Ueda, Tetsuya
TITLE OF INVENTION: TRANSCRIPTION FACTOR REGULATING FGF-2
FILE REFERENCE: SALKINS.015A
CURRENT APPLICATION NUMBER: US/09/079,431B
CURRENT FILING DATE: 1998-05-14
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 630
TYPE: PRT
ORGANISM: Homo sapiens
US-09-079-431B-2

Query Match 9.9%; Score 214.5; DB 4; Length 630;
Best Local Similarity 45.7%; Pred. No. 5.2e-09;
Matches 42; Conservative 13; Mismatches 30; Indels 7; Gaps 1;

QY 150 MDCPALPGMKKEVIRKSGLSAGKSDVYFSPSGKFRSKPOLARYIGNTVDSLSPDR 209
DB 6 LDCPALPGMKKEVIRKSGATCGRSPTYOSPTGDIRSKVELTRYLGPACDITLDFDK 65
QY 210 TG-----KMPKSKLOKNKORLNDPLNOK 234
DB 66 OGILCYPAKAPHPVAVASKKKRKPAPKTRK 97

RESULT 5

US-09-249-585A-3
Sequence 3, Application US/09249585A
Patent No. 6417002
GENERAL INFORMATION:
APPLICANT: Horlick, Robert
TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES
FILE REFERENCE: 0867/0D905
CURRENT APPLICATION NUMBER: US/09/249,585A
CURRENT FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 641
TYPE: PRT

Db 125 -----GGGGGGGGG 134

RESULT 10
US-09-478-948-6
Sequence 6, Application US/09478948
Patent No. 6465258

GENERAL INFORMATION:

APPLICANT: Shan, Bel

APPLICANT: Okamoto, Arthur Y.

APPLICANT: Tularik Inc.

TITLE OF INVENTION: FXR Receptor-Mediated Modulation of Cholesterol

FILE REFERENCE: 018781-00131005

CURRENT APPLICATION NUMBER: US/09/478,948

CURRENT FILING DATE: 2000-01-06

PRIOR FILING DATE: 1999-01-07

NUMBER OF SEQ ID NOS: 6

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 6

LENGTH: 201

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: flexible linker

NAME/KEY: MOD_RES

LOCATION: (1)-(97)

OTHER INFORMATION: Gly at positions 1-97 may be present or absent

NAME/KEY: MOD_RES

LOCATION: (105)-(201)

OTHER INFORMATION: Gly at positions 105-201 may be present or absent

US-09-478-948-6

Query Match 9.7%; Score 209.5; DB 4; Length 201;
Best local Similarity 40.7%; Pred. No. 2.9e-09;
Matches 55; Conservative 0; Mismatches 63; Indels 17; Gaps 2;

QY 6 GGGHCCPEDEEGESAGSGAGDSALIEQGGGSSALAPSPVSGVRRREGARGGGRGRMK 65
Db 17 GGG-----GG 68
QY 66 GAGGGGCGVCGRGRCRGRCRGRCRGRCRGRCRGRCRGRCRGRCRGRCRGRCRGRCRG 125
Db 69 GGG 124
QY 126 PVPPSGSAGCPGPG 140
Db 125 -----GGGGGGGGGG 134

RESULT 11
US-08-806-029-9
Sequence 9, Application US/08806029
Patent No. 6380154

GENERAL INFORMATION:

APPLICANT: Cappello, Joseph

APPLICANT: Stedronsky, Erwin R.

TITLE OF INVENTION: Synthetic Proteins for in vivo Drug

TITLE OF INVENTION: Delivery and Tissue Augmentation

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: United States

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/806,029

FILING DATE: 24-FEB-1997

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,237

FILING DATE: 11-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-58847-2/REF/MTK

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

INFORMATION FOR SEQ ID NO. 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1136 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-806-029-9

Query Match 9.1%; Score 196.5; DB 4; Length 1136;
Best local Similarity 33.0%; Pred. No. 2.9e-07;
Matches 59; Conservative 10; Mismatches 95; Indels 15; Gaps 5;

QY 3 AHGGGRCCEDEEGESAGSGAGDSALIEQ-GGGSSALAPSPVSGVRRREGARGGGRGR 61
Db 965 AGAGSGAGAGSGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1024
QY 62 GRWKQAGRGCGVCGRGRCRGRCRGRCRGRCRGRCRGRCRGRCRGRCRGRCRGRCRG 121
Db 1025 GAGSAG 1083
QY 122 PRPPVFPSSGSGPGR 180
Db 1084 GAGSAG 1129

RESULT 12
US-07-609-716-31
Sequence 31, Application US/07609716
Patent No. 551581

GENERAL INFORMATION:

APPLICANT: Ferrari, Franco A.

APPLICANT: Cappello, Joseph

TITLE OF INVENTION: Functional Recombinantly Prepared

TITLE OF INVENTION: Synthetic Protein Polymer

NUMBER OF SEQUENCES: 118

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: US

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/609,716

FILING DATE: 06-NOV-1990

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rowland, Bertam I

REGISTRATION NUMBER: 20015

REFERENCE/DOCKET NUMBER: A-55186-3/BIR

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

;; INFORMATION FOR SEQ ID NO: 31:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1177 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-07-609-716-31

Query Match 9.1%; Score 196.5; DB 1; Length 1177;
Best Local Similarity 33.0%; Pred. No. 3.1e-07;
Matches 59; Conservative 10; Mismatches 95; Indels 15; Gaps 5;

QY 3 AHFGGRCCEPQEEGESASAGSGAGSDAI EOG-GGGSALAPSPVSVRRREGARCGGR 61
DB 1006 AGAGSGAGAGSGAAGYGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGA 1065
QY 62 GRWKQARGGVCGRCGRGRGRGRGRGRGRPPSGSGLGDDGGCGGGGGGGA 121
DB 1066 GAGSGAGAGSGAAGYGAGAGSGAGAGSGAGAGSGAGAGSGAG-AGSGAGAGSGAGAGSGA 1124
QY 122 PRREPVPFSGSAGPGRGRPRATESGKRMDCPALPGMKKEEVIRKSGLSAGSDVYF 180
DB 1125 GAGSGAGAGSGAAGYG--AGAGSGAGAGSGAGAMPD-----RYQLSAGR---YHY 1170

RESULT 13

US-08-175-155-29
; Sequence 29, Application US/08175155
; Patent No. 5641648

;; GENERAL INFORMATION:
;; APPLICANT: Ferrari, Franco A.
;; APPLICANT: Cappello, Joseph
;; APPLICANT: Crissman, John W.
;; APPLICANT: Dorman, Mary A.
;; TITLE OF INVENTION: Methods for Preparing Synthetic
;; TITLE OF INVENTION: Repetitive DNA
;; NUMBER OF SEQUENCES: 69
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;; STREET: Four Embarcadero Center, Suite 3400
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: US
;; ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,155
FILING DATE: 29-DEC-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Howland, Berttram I.
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: A-55186-5/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 29:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1177 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-175-155-29

Query Match 9.1%; Score 196.5; DB 1; Length 1177;
Best Local Similarity 33.0%; Pred. No. 3.1e-07;
Matches 59; Conservative 10; Mismatches 95; Indels 15; Gaps 5;

QY 3 AHFGGRCCEPQEEGESASAGSGAGSDAI EOG-GGGSALAPSPVSVRRREGARCGGR 61
DB 1006 AGAGSGAGAGSGAAGYGAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGA 1065
QY 62 GRWKQARGGVCGRCGRGRGRGRGRGRGRPPSGSGLGDDGGCGGGGGGGA 121
DB 1066 GAGSGAGAGSGAAGYGAGAGSGAGAGSGAGAGSGAGAGSGAG-AGSGAGAGSGAGAGSGA 1124
QY 122 PRREPVPFSGSAGPGRGRPRATESGKRMDCPALPGMKKEEVIRKSGLSAGSDVYF 180
DB 1125 GAGSGAGAGSGAAGYG--AGAGSGAGAGSGAGAMPD-----RYQLSAGR---YHY 1170

RESULT 14

US-08-477-509B-64
; Sequence 64, Application US/08477509B
; Patent No. 5770697

;; GENERAL INFORMATION:
;; APPLICANT: Ferrari, Franco A.
;; APPLICANT: Cappello, Joseph
;; APPLICANT: Crissman, John W.
;; APPLICANT: Dorman, Mary A.
;; TITLE OF INVENTION: No. 5770697e1 Peptides Comprising Repetitive
;; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
;; NUMBER OF SEQUENCES: 112
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;; STREET: Four Embarcadero Center, Suite 3400
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: US
;; ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,509B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,155
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,049
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-7/RFT/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 64:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1177 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-477-509B-64

Query Match 9.1%; Score 196.5; DB 1; Length 1177;
Best Local Similarity 33.0%; Pred. No. 3.1e-07;
Matches 59; Conservative 10; Mismatches 95; Indels 15; Gaps 5;

XX
PS Disclosure; Fig 9c; 114pp; English.

CC This sequence is the human DNA demethylase, designated dmtasel, of
CC the invention. The DNA demethylase is overexpressed in cancer cells.
CC Expression of the demethylase cDNA is useful to alter DNA methylation
CC patterns of DNA in vitro in cells or in vivo in humans, animals and
CC plants. The cDNA is in antisense orientation to inhibit demethylase in
CC cancer cells for therapeutic purposes. The demethylase is used to alter
CC the differentiation state and to generate stem cells for therapeutics.
CC The differentiation state and to improve expression of foreign genes. The
CC cDNA can also be used for recombinant production of large amounts of the
CC demethylase. The protein can be used for high throughput screening of
CC demethylase inhibitors, and for obtaining the x-ray crystal structure.
CC The demethylase cDNA and protein are also useful for changing the state
CC of differentiation of a cell to allow gene therapy, stem cell selection
CC or cell cloning; or for inhibiting methylation in cancer cells using
CC vector mediated gene therapy. Antagonists or inhibitors of the
CC demethylase can be used to manufacture medicaments for cancer treatment,
CC for restoring an aberrant methylation patterns or changing methylation
CC patterns in patient DNA. Change of the methylation pattern activates a
CC silent gene permitting the correction of a genetic defect, such as a
CC beta-thalassemia or sickle cell anemia. The cDNA can be used as a
CC template to design antisense oligonucleotides and ribozymes. The cDNA can
CC also be used in two-hybrid systems in yeast to identify proteins
CC interacting with demethylase. Determining the level of DNA methylation
CC expression in a cell can be used as a method for diagnosis of cancer,
CC where overexpression is indicative of cancer cells.

XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 2167; DB 20; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.1e-165;

Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAHPGGGRCPEDEBESNAAGSGAGDSAIIEGGGGSALAPSPVSGVREGARGGARG 60
DB 1 MRAHPGGGRCPEDEBESNAAGSGAGDSAIIEGGGGSALAPSPVSGVREGARGGARG 60
QY 61 RGRWKQAGRGGVCGR 120
DB 61 RGRWKQAGRGGVCGR 120
QY 121 APRREPVPFPGSGAGPGRPRATESGKRMDCPALPPEWKKEEYIRKSGLSAGKSDYYF 180
DB 121 APRREPVPFPGSGAGPGRPRATESGKRMDCPALPPEWKKEEYIRKSGLSAGKSDYYF 180
QY 181 SPSGKFRSKPOLARYLGNVTDLSSFDPRGKMPKSLQKNKORLNDPLNKNKGRPDNL 240
DB 181 SPSGKFRSKPOLARYLGNVTDLSSFDPRGKMPKSLQKNKORLNDPLNKNKGRPDNL 240
QY 241 TTLPIRGTASTFKOPVTYVTHNPSKNKVASDPQRMNEQPROLFWEKRLQGLSASVTBII 300
DB 241 TTLPIRGTASTFKOPVTYVTHNPSKNKVASDPQRMNEQPROLFWEKRLQGLSASVTBII 300
QY 301 KTMELPKGLGGGSGNSNETLISAVASALHTSSAPITGOVSAAEKNPAVWLNTSOPPLCK 360
DB 301 KTMELPKGLGGGSGNSNETLISAVASALHTSSAPITGOVSAAEKNPAVWLNTSOPPLCK 360
QY 361 AFIVTDEDIRKQEEERVQOVRRKLEBALMADILSRAADTEEMDIEMDSDEA 411
DB 361 AFIVTDEDIRKQEEERVQOVRRKLEBALMADILSRAADTEEMDIEMDSDEA 411

RESULT 2
ID AAB99915
AA AAB9915 standard; Protein: 411 AA.

XX AAB9915;

XX 26-SEP-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:1.

XX Differentiation; heart muscle cell; cytokine; transcription factor;
KW proliferation; surface antigen; heart disease; cardiomyocyte;
KW bone marrow; umbilical blood cell; heart muscle degeneration;
KW myocardial infarction.

XX Homo sapiens.

XX WO200148150-A1.

XX 05-JUL-2001.

XX 02-NOV-2000; 2000WO-JP07741.

XX 28-DEC-1999; 99JP-0372826.

XX 28-FEB-2000; 2000WO-JP01148.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;
PI Yamada Y;

XX WPI: 2001-425655/45.

XX N-PSDB: AAH44351.

XX Cells capable of differentiating into cardiomyocytes and originating in
PT bone marrow or umbilical blood cells for study of cardiomyocyte
PT differentiation and treatment of heart disease

XX Claim 22; Page 84-86; 187pp; Japanese.

CC The present invention describes cells originating in bone marrow or
CC umbilical blood cells which are capable of differentiating into
CC cardiomyocytes. Also described are: (1) cardiomyocytes produced by the
CC differentiation into cardiomyocytes, regulated by a promotional and/or
CC inhibitory factor; (2) a method for the differentiation of the cells
CC into cell types other than cardiomyocytes; (3) drug compositions
CC promoting the formation of heart muscle and regeneration of heart tissue
CC which contain the cells; (4) a method for the production of heart tissue
CC which recognise the cells; (5) a method for screening factors which
CC promote the proliferation of the cells; (6) a method for screening factors which
CC promote the proliferation of the cells; (7) a method for immortalising
CC the cells by expressing telomerase in them; (8) drug compositions for
CC the treatment of heart disease which contain the immortalised cells; and
CC (9) cell-free supernatant from the culture of the cells and its use in
CC promoting their differentiation into cardiomyocytes. The cells are used
CC in the treatment of diseases involving heart muscle degeneration, such
CC as myocardial infarction and in the study of cardiomyocyte
CC differentiation. AAH44351 to AAH44409 and AAB9915 to AAB99915 represent
CC sequences used in the exemplification of the present invention.

XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 2167; DB 22; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.1e-165;

Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAHPGGGRCPEDEBESNAAGSGAGDSAIIEGGGGSALAPSPVSGVREGARGGARG 60
DB 1 MRAHPGGGRCPEDEBESNAAGSGAGDSAIIEGGGGSALAPSPVSGVREGARGGARG 60
QY 61 RGRWKQAGRGGVCGR 120
DB 61 RGRWKQAGRGGVCGR 120
QY 121 APRREPVPFPGSGAGPGRPRATESGKRMDCPALPPEWKKEEYIRKSGLSAGKSDYYF 180
DB 121 APRREPVPFPGSGAGPGRPRATESGKRMDCPALPPEWKKEEYIRKSGLSAGKSDYYF 180
QY 181 SPSGKFRSKPOLARYLGNVTDLSSFDPRGKMPKSLQKNKORLNDPLNKNKGRPDNL 240
DB 181 SPSGKFRSKPOLARYLGNVTDLSSFDPRGKMPKSLQKNKORLNDPLNKNKGRPDNL 240

XX	PA	(META-) METAGEN GES GENOMFORSCHUNG MBH.
XX	PI	Rosenthal A, Specht T, Hinzmann B, Schmittl A, Pilarsky C, Dahl E;
XX	DR	WPI; 1999-621386/54.
XX	DR	N-PSDB; AA52863.
XX	PT	New human nucleic acid sequences from pancreatic tumors, and related proteins -
XX	PS	ClaIm 23; Page 315; 502pp; German.
XX	CC	This invention describes novel polypeptides and their encoding nucleic acid derived from human pancreatic tumor tissue which have cytostatic activity. The sequences are also useful in producing pharmaceutical compositions for treatment of pancreatic tumors. AAY73814-Y74252
XX	CC	represent protein fragments encoded by the human pancreatic tumor cDNA library derived expressed sequence tag (EST) sequences represented in
XX	CC	AA52858-253014.
SQ	Sequence	281 AA;
Query Match	Best Local Similarity	66.7%; Score 1445; DB 20; Length 281;
Matches 281; Conservative % 0; Mismatches 0; Indels 0; Gaps 0;		
OY	131	SGSAGCPRRPRATEGSKRMDCPALPFGKKKEEYTRKSGLSAGKSDVYYSPSGKKRFSK 190 Db 1 SGSAGCPRRPRATEGSKRMDCPALPFGKKKEEYTRKSGLSAGKSDVYYSPSGKKRFSK 60
OY	191	POLARYLGNTVDLSEDFRTGKMMSKLQKNQRRLANDPLNQNGKDPDLMTLPPIROTAS 250 Db POLARYLGNTVDLSEDFRTGKMMSKLQKNQRRLANDPLNQNGKDPDLMTLPPIROTAS 120
OY	251	IFKPOTTKTYTNHPSNKKVKSDPORMNEOPROLFEWEKRLOGLSASDVTEQQIITKTMELPKGIQ 310 Db IFKPOTTKTYTNHPSNKKVKSDPORMNEOPROLFEWEKRLOGLSASDVTEQQIITKTMELPKGIQ 180
OY	311	GVPGSNDTELFLSAVASALHTSSAPLTGOVSAAVEKNPAVWLNTSOPLCAFIYVTDEDIR 370 Db GVPGSNDTELFLSAVASALHTSSAPLTGOVSAAVEKNPAVWLNTSOPLCAFIYVTDEDIR 240
OY	371	KOERVOOVRRKIIEELMADILSRADTPEMDIEMDSGEA 411 Db KOERVOOVRRKIIEELMADILSRADTPEMDIEMDSGEA 281
RESULT 7		
AAY48439	ID	AAY48439 standard; Protein; 281 AA.
AC	AAV48439;	
DT	08-DEC-1999	(first entry)
DE	Human prostate cancer-associated protein 136.	
KW	Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;	
KW	cancer; tissue specificity; human.	
OS	Homo sapiens.	
PN	DEL19811194-A1.	
PD	16-SEP-1999.	
PF	10-MAR-1998; 98DE-1011194.	
PR	10-MAR-1998; 98DE-1011194.	
PA	(META-) METAGEN GES GENOMFORSCHUNG MBH.	

PI	Specht T,	Hinzmann B,	Schmitt A,	Pilarsky C,	Dahl E,	Rosenthal A;
XX						
DR	WPI: 1999-519629/44.					
DR	N-PSDB: AA233533.					
XX						
PT	New nucleic acid expressed at high level in normal prostatic tissue and					
PT	encoded polypeptides, used to treat cancer and screen for therapeutic					
XX	agents					
PS	Claim 22: 176; 194pp; German.					
XX						
CC	This invention describes novel nucleic acid sequences (A) that are					
CC	expressed at high level in normal prostatic tissue. Polypeptides (I)					
CC	encoded by (A) are used: (a) for identifying agents for treatment of					
CC	prostatic cancer and (b) for therapy of prostate cancer, optionally					
CC	where expressed by gene therapy methods. (A) is also used to isolate					
CC	full-length genes (for gene therapy) and for recombinant production of					
CC	(I), which can be used to raise specific antibodies. (A) are identified					
CC	by assembly of ESTs (expressed sequence tags) before these are analyzed					
CC	for expression pattern (tissue specificity). This approach eliminates					
CC	many of the false results, as regards tissue specificity, associated					
CC	with known methods that use single (usually short) ESTs. AA948304-V484555					
CC	represent peptides encoded by the expressed sequence tags described in					
CC	the method of the invention.					
XX						
Sequence	281 AA;					
XX						

[illegible]


```

OY 210 TGKMPSTLQKKNKORLNDPLNOKKGPDLNTTLPPIROTASTFKQPTKVTNHPNSNKVKS 269
DB 61 TGKMPSTLQKKNKORLNDPLNOKKGPDLNTTLPPIROTASTFKQPTKVTNHPNSNKVKS 120
OY 270 DQQRNEQPPQLEFWEKRLQGLSASDVTEQIITKTMELPKGLQGVGSGNDETLLSAVASAL 329
DB 121 DQQRNEQPPQLEFWEKRLQGLSASDVTEQIITKTMELPKGLQGVGSGNDETLLSAVASAL 180
OY 330 HTSSAPITGVSAVAEKNPAYWLNTSOPLCARFIVTDEDIRKQBERVQVRRKLEALMA 389
DB 181 HTSSAPITGVSAVAEKNPAYWLNTSOPLCARFIVTDEDIRKQBERVQVRRKLEALMA 240
OY 390 DILSRADTEEMDIEMDSGDEA 411
DB 241 DILSRADTEEMDIEMDSGDEA 262

RESULT 9
AAE22578
ID AAE22578 standard; Protein: 282 AA.
AC AAE22578;
XX
DT 26-JUL-2002 (first entry)
XX
DE Xenopus laevis-MBD3 protein.
XX
KW Gene expression; cellular chromatin; methyl CpG binding domain; cancer;
KW localisation domain; diabetic retinopathy; ischaemia; HIV infection;
KW human immuno deficiency virus; macular degeneration; vascular disease;
KW rheumatoid arthritis; psoriasis; Alzheimer's disease; muscular dystrophy;
KW sickle cell anaemia; stroke; neurodegenerative disease; cystic fibrosis;
KW gene therapy; cytostatic; antidiabetic; ophthalmological; vasotrophic;
KW neuroprotective; nontropic; cerebroprotective; antibacterial; antifungal;
KW antiviral; MBD3 protein.
XX
OS Xenopus laevis.
XX
FH Key 1..69 Location/Qualifiers
FT Binding-site /label= Methyl_CpG_binding_domain
FT Region 7..9
FT /label= Beta1_helix
FT /label= Beta2_helix
FT /label= Beta2_helix
FT /label= L1_loop
FT /label= Beta3_helix
FT /label= Beta3_helix
FT /label= Beta4_helix
FT /label= Alpha1_helix
FT /label= L2_loop
FT /label= L2_loop
FT /label= Hairpin_loop
FT Region 61..68
FT /label=
XX
XX WO200226960-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US42377.
XX
XX 29-SEP-2000; 2000US-236884P.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX Wolfe AP, Urnov F, Lai A, Raschke E;
XX
XX MPI; 2002-372124/40.
XX

```

```

PT Compartmentalising a region of interest in cellular chromatin which
PT facilitates the modulation of the expression of a gene comprises
PT contacting the gene with a composition comprising a localization domain
PT and a DNA binding domain -
XX
XX Example 2; Fig 1B; 85pp; English.
XX
XX The present invention relates to methods and compositions for regulating
XX gene expression. In particular the method of compartmentalising a region
XX of interest in cellular chromatin comprises contacting the region of
XX interest with a composition that binds to a binding site in cellular
XX chromatin, where the binding site is in a gene of interest and the
XX composition comprises a localisation domain (e.g., methyl CpG binding
XX domain obtained from MBD1, MBD2, MBD3, dMBD-1like and dMBD-1like
XX delta) and a DNA binding domain (or functional fragment). The method is
XX useful for compartmentalising a region of interest in cellular chromatin
XX which facilitates the modulation of the expression of a gene using a
XX fusion molecule comprising a DNA binding domain and a localisation domain
XX that binds to the chromatin. The fusion molecules or polypeptides can be
XX used to prepare pharmaceutical compositions to prevent or treat cancer,
XX ischaemia, diabetic retinopathy, macular degeneration, HIV infection,
XX rheumatoid arthritis, psoriasis, sickle cell anaemia, vascular disease,
XX Alzheimer's disease, muscular dystrophy, neurodegenerative diseases,
XX cystic fibrosis, stroke, bacterial, viral or fungal infections. Sequences
XX of the invention are also used in gene therapy. The present sequence is
XX Xenopus laevis MBD3 LF protein. This sequence is used in the
XX exemplification of the invention.
XX
XX Sequence 282 AA:
XX
XX Query Match 48 98; Score 1059; DB 23; Length 282;
XX Best Local Similarity 75.48; Pred. No. 6; 5e-77;
XX Matches 202; Conservative 33; Mismatches 27; Indels 6; Gaps 2;
XX
OY 148 KRMDCPALPPGKKKEVIRKSGISAGSDVYFSPSGKKFRSKPOLARLYLGNVTLDSSPD 207
DB 4 KRMECSAL-QGKKKEVTRSRGSLGSAKSDVYFSPSGKKFRSKPOLARLYLGNVTLDSSPD 62
OY 208 FRFGKMPSTLQKKNKORLNDPLNOKKGPDLNTTLPPIROTASTFKQPTKVTNHPNSNKV 267
DB 63 FRFGKMPSTLQKKNKORLNDPLNOKKGPDLNTTLPPIROTASTFKQPTKVTNHPNSNKV 122
OY 268 KSDPQRNEQPPQLEFWEKRLQGLSASDVTEQIITKTMELPKGLQGVGSGNDETLLSAVAS 327
DB 123 KSDPQRNEQPPQLEFWEKRLQGLSASDVTEQIITKTMELPKGLQGVGSGNDETLLSAVAS 182
OY 328 ALHTSSAPITGVSAVAEKNPAYWLNTSOPLCARFIVTDEDIRKQBERVQVRRKLEAL 387
DB 183 ALHTSSAPITGVSAVAEKNPAYWLNTSOPLCARFIVTDEDIRKQBERVQVRRKLEAL 242
OY 388 MADILSRADTEEMDIEMDSGDE 410
DB 243 MADILSRADTEEMDIEMDSGDE 270

RESULT 10
AAE22577
ID AAE22577 standard; Protein: 303 AA.
AC AAE22577;
XX
XX 26-JUL-2002 (first entry)
XX
XX Xenopus laevis MBD3 LF protein.
XX
XX Gene expression; cellular chromatin; methyl CpG binding domain; cancer;
XX localisation domain; diabetic retinopathy; ischaemia; HIV infection;
XX human immuno deficiency virus; macular degeneration; vascular disease;
XX rheumatoid arthritis; psoriasis; Alzheimer's disease; muscular dystrophy;
XX sickle cell anaemia; stroke; neurodegenerative disease; cystic fibrosis;
XX gene therapy; cytostatic; antidiabetic; ophthalmological; vasotrophic;
XX neuroprotective; nontropic; cerebroprotective; antibacterial; antifungal;
XX antiviral; MBD3 protein.
XX

```

```

XX OS Xenopus laevis.
XX
FH Key Location/Qualifiers
FT Binding-site 1..90
FT /label= Methyl_CpG_binding_domain
FT Region 7..9
FT /label= Beta1_helix
FT Region 15..21
FT /label= Beta2_helix
FT Region 22..31
FT /label= L1_loop
FT Region 32..37
FT /label= Beta3_helix
FT Region 61..63
FT /label= Beta4_helix
FT Region 67..73
FT /label= Alpha1_helix
FT Region 74..80
FT /label= L2_loop
FT Region 82..89
FT /label= Hairpin_loop
XX
XX WO200226960-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US42377.
XX
XX 29-SEP-2000; 2000US-236884P.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX Wolfe AP, Urnov F, Lai A, Raschke E;
XX
XX WPI: 2002-372124/40.
XX
XX
XX Compartmentalizing a region of interest in cellular chromatin which
XX facilitates the modulation of the expression of a gene comprises
XX contacting the gene with a composition comprising a localization domain
XX and a DNA binding domain -
XX
XX Example 2; Fig 1B; 85pp; English.
XX
XX The present invention relates to methods and compositions for regulating
XX gene expression. In particular the method of compartmentalizing a region
XX of interest in cellular chromatin comprises contacting the region of
XX interest with a composition that binds to a binding site in cellular
XX chromatin, where the binding site is in a gene of interest and the
XX composition comprises a localization domain (e.g., methyl CpG binding
XX domain obtained from MBDP2, MBD1, MBD2, MBD3, dMBD-like and dMBD-like
XX delta) and a DNA binding domain (or functional fragment). The method is
XX useful for compartmentalizing a region of interest in cellular chromatin
XX which facilitates the modulation of the expression of a gene using a
XX fusion molecule comprising a DNA binding domain and a localization domain
XX that binds to the chromatin. The fusion molecules or polypeptides can be
XX used to prepare pharmaceutical compositions to prevent or treat cancer,
XX ischaemia, diabetic retinopathy, macular degeneration, HIV infection,
XX rheumatoid arthritis, psoriasis, sickle cell anaemia, vascular disease,
XX Alzheimer's disease, muscular dystrophy, neurodegenerative diseases,
XX cystic fibrosis, stroke, bacterial, viral or fungal infections. Sequences
XX of the invention are also used in gene therapy. The present sequence is
XX exemplification of the invention.
XX
XX Sequence 303 AA:
XX
XX Query Match 48.8%, Score 1056.5; DB 23; Length 303;
XX Best Local Similarity 70.5%; Pred. No. 1.le-76;
XX Matches 203; Conservative 33; Mismatches 27; Indels 25; Gaps 2;
XX
XX 148 KMDPAPLPQWKKKEVTRKSLGKSDVYFS-----PSGKFF 187
XX | : | | | | | | | | | | | | | | | | | | | |

```

```

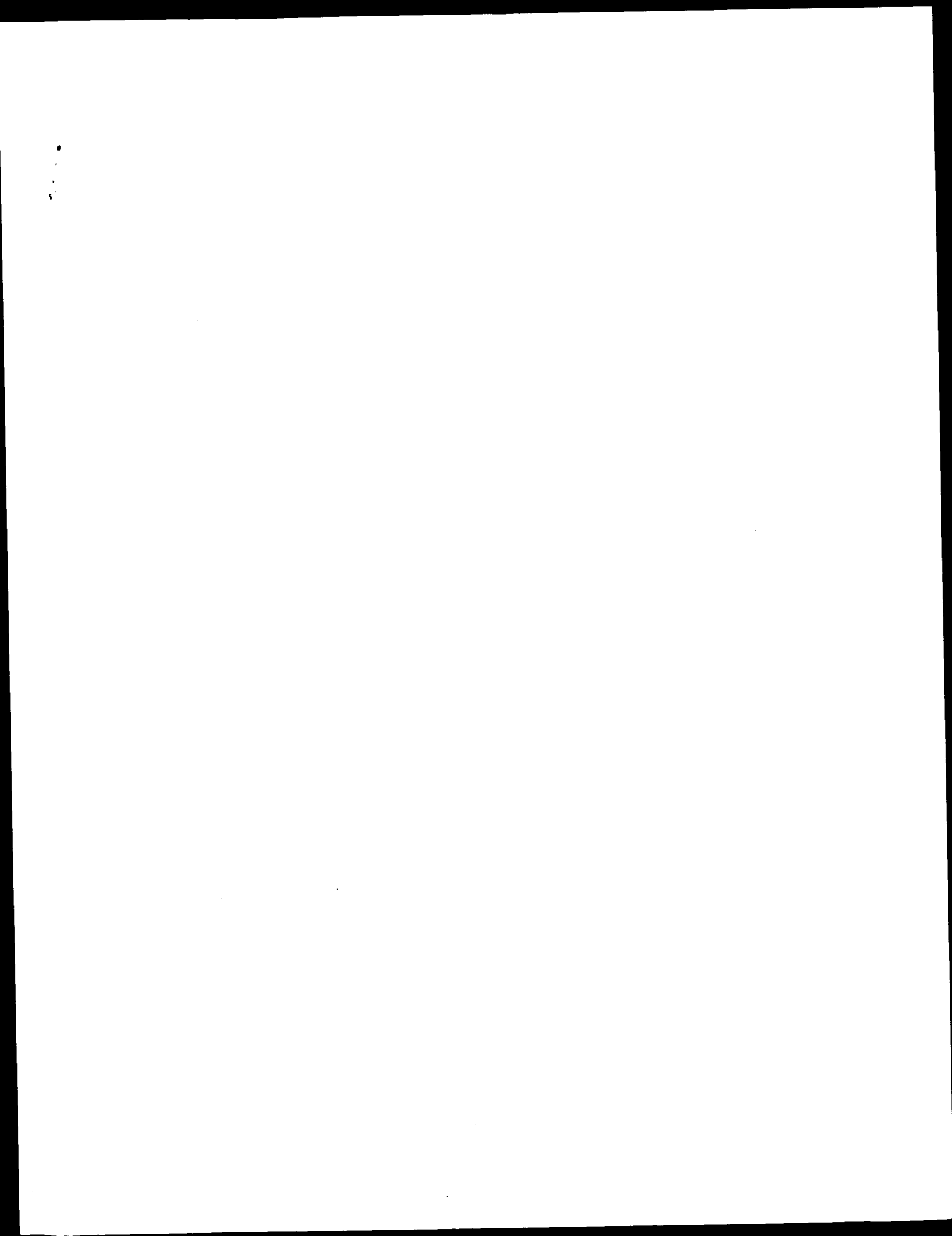
DB 4 KRWCSSALPQGWKKKEVTRRSGLSAGKSDVYYSPPSRNRSIDRVGCLINPSSGKF 63
OY 188 RSKPOLARLYGNTVLSFDFPTGKMPKSKLOKNNORLRNDPLNONGKPLDNTLPTRQ 247
DB 64 RSKPOLARLYGNSMDLSTFDFPTGKMLSKINKNRRKRYGLNOSKPKPLNTRALPYRQ 123
OY 248 TASIFKOPVTKYTNPSNKKVKSDDPQRMNDOPRQLFWERKLOGLSADVTEOIIKTMELPK 307
DB 124 TASIFKOPVTKYTNPTNKKVKSDDPQKAVDQPRQLFWERKLSGLNAPDIAEELVYTMELPK 183
OY 308 GLQGVGPGSNDTELLSAVASALHTSSAPITGOVSAVKNRPAWLNTPQPLCKAFIYVDE 367
DB 184 GLQGVGPGCTDETLTSAIASALHTSTMTPTGOLSAVEKNPQVWLNTPQPLCKAFWYVDE 243
OY 368 DIRKQERYQVYRKKEEFALMDILSRADTPE-----MDTMDSGDE 410
DB 244 DIRKQERYQVYRKKEEFALMDILAHVEISKDGGAFLDIDDEE 291

RESULT 11
AA14198
ID AA14198 standard; Protein; 291 AA.
XX
XX AA14198;
XX
XX 28-JUL-1999 (first entry)
XX
XX Human DNA demethylase, dmtase2, protien sequence.
XX
XX DNA demethylase; dmtase1; dmtase2; human; cancer cell; beta-thalassemia;
XX DNA methylation pattern; methylation inhibitor; gene therapy; diagnosis;
XX genetic defect correction; sickle cell anaemia.
XX
XX Homo sapiens.
XX
XX WO9924583-A1.
XX
XX 20-MAY-1999.
XX
XX 12-NOV-1998; 98WO-CA01059.
XX
XX 11-MAY-1998; 98CA-2230991.
XX
XX 12-NOV-1997; 97CA-2220805.
XX
XX (UYMC-) UNIV MCGILL.
XX
XX Bhatnagar S, Ramchandani S, Szyf M;
XX
XX WPI: 1999-347283/29.
XX
XX N-PSDB; AAX61219.
XX
XX Human and murine DNA demethylases useful in the diagnosis of cancer
XX
XX Disclosure: Fig 9f; 114pp; English.
XX
XX This sequence is the human DNA demethylase, designated dmtase2, of
XX the invention. The DNA demethylase cDNA is overexpressed in cancer cells.
XX Expression of DNA demethylase cDNA is useful to alter DNA methylation
XX patterns of DNA in vitro in cells or in vivo in humans, animals and
XX plants. The cDNA is in antisense orientation to inhibit demethylase in
XX cancer cells for therapeutic purposes. The demethylase is used to alter
XX the differentiation state and to generate stem cells for therapeutics,
XX cells for animal cloning and to improve expression of foreign genes. The
XX cDNA can also be used for recombinant production of large amounts of the
XX demethylase. The protein can be used to raise antibodies against
XX demethylase inhibitors, and for obtaining the x-ray crystal structure.
XX The demethylase cDNA and protein are also useful for changing the state
XX of differentiation of a cell to allow gene therapy, stem cell selection
XX or cell cloning; or for inhibiting methylation in cancer cells using
XX vector mediated gene therapy. Antagonists or inhibitors of the
XX demethylase can be used to manufacture medicaments for cancer treatment,
XX for restoring an aberrant methylation patterns or changing methylation

```


PF 15-JUL-1998; 98WO-US14679.
 XX
 PR 22-JUN-1998; 98US-0102322.
 PR 17-JUL-1997; 97US-0896164.
 PR 10-OCT-1997; 97US-0061599.
 PR 10-OCT-1997; 97US-0061765.
 PR 10-OCT-1997; 97US-0948705.
 PR 11-OCT-1997; 97GB-0021697.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Chen Y, Gout I, Gure A, O'Hare M, Ohta Y, Old LJ;
 PI Pfeundschnh M, Sahin U, Scanlan MJ, Stockert E;
 PI Tureci O;
 XX
 DR WPI: 1999-132448/11.
 XX
 PT New isolated cancer associated nucleic acids and polypeptides -
 PT isolated using sera from cancer patients, used to develop products
 PT for the diagnosis, monitoring or treatment of cancers
 XX
 PS Disclosure: Page 677; 787pp; English.
 XX
 CC The invention relates to a method for diagnosing a disorder characterised
 CC by expression of a human cancer associated antigen precursor coded for by
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
 CC biological sample isolated from a subject with an agent that specifically
 CC binds to the NAM, an expression product or a fragment of an expression
 CC product complexed with an HLA molecule; and (b) determining the
 CC interaction between the agent and the NAM or the expression product as a
 CC determination of the disorder. The products and methods can be used in
 CC the diagnosis, monitoring, research, or treatment of conditions
 CC characterised by the expression of various cancer associated antigens.
 CC The invention provides nucleic acid sequences and encoded polypeptides
 CC which are cancer associated antigen precursors expressed in human breast
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC lung cancer.
 CC
 XX
 SQ Sequence 200 AA;
 Query Match 46.7%; Score 1013; DB 20; Length 200;
 Best Local Similarity 100.0%; Pred. No. 2.1e-73;
 Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 212 KMPSKLOKNNKRLRNDPLNCKGKPDINTLPIRQTASIFKQPVTKTNHPSKVKSDP 271
 DB 1 KMPSKLOKNNKRLRNDPLNCKGKPDINTLPIRQTASIFKQPVTKTNHPSKVKSDP 60
 QY 272 QRANQPRQLFWERKLGSLASDVTEQIITKTMELPKGLQGVGSGSNDFTLLSAVALHT 331
 DB 61 QRANQPRQLFWERKLGSLASDVTEQIITKTMELPKGLQGVGSGSNDFTLLSAVALHT 120
 QY 332 SSAPITGVSAAVEKNPAVWLNTPSOPICAFIYDDEDIRKQEEVQVQRKKLEALMADI 391
 DB 121 SSAPITGVSAAVEKNPAVWLNTPSOPICAFIYDDEDIRKQEEVQVQRKKLEALMADI 180
 QY 392 LSRADTFEEMDIEMDSGEA 411
 DB 181 LSRADTFEEMDIEMDSGEA 200
 RESULT 14
 AAB92997
 ID AAB92997 standard; Protein: 223 AA.
 XX
 AC AAB92997;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:11731.
 XX
 KM Human: primer: detection; diagnosis; antisense therapy; gene therapy.
 XX

OS Homo sapiens.
 XX
 PN EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 11731; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 CC
 XX
 SQ Sequence 223 AA;
 Query Match 34.7%; Score 753; DB 22; Length 223;
 Best Local Similarity 70.5%; Pred. No. 1.8e-52;
 Matches 148; Conservative 25; Mismatches 25; Indels 12; Gaps 2;
 QY 213 KMPSKLOKNNKRLRNDPLNCKGKPDINTLPIRQTASIFKQPVTKTNHPSKVKSDP 272
 DB 1 KMPSKLOKNNKRLRNDPLNCKGKPDINTLPIRQTASIFKQPVTKTNHPSKVKSDP 60
 QY 273 RMNEQPRQLFWERKLGSLASDVTEQIITKTMELPKGLQGVGSGSNDFTLLSAVALHTS 332
 DB 61 KAVDQPRQLFWERKLGSLASDVTEQIITKTMELPKGLQGVGSGCIDEFTLLSAVALHTS 120
 QY 333 SAPITGVSAAVEKNPAVWLNTPSOPICAFIYDDEDIRKQEEVQVQRKKLEALMADI 392
 DB 121 TWPITGOLSAAVEKNPAVWLNTPSOPICAFIYDDEDIRKQEEVQVQRKKLEALMADI 180
 QY 393 SR-----ADTE-----EMDIEMDSGE 410
 DB 181 AHVEELARDGEAPLDKACAEDEDEDEEE 210



GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2003, 03:11:54 ; Search time 10.511 Seconds

(without alignments)
1033.836 Million cell updates/sec

Title: US-09-554-414b-2_COPY_150_411

Perfect score: 1344

Sequence: 1 MDCPALPPGKKKEVIRKSG.....LSRAADTEMDIEMDSGDEA 262

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	177	13.2	486	MEC2_HUMAN	P51608 homo sapien
2	167	12.4	484	MEC2_MOUSE	O92206 mus musculus
3	167	12.4	492	MEC2_RAT	O00566 rattus norv
4	107.5	8.0	1039	MSL1_DROME	P50535 drosophila
5	97	7.2	448	YAP1_CHICK	P46936 gallus gall
6	95	7.1	612	ADP1_CANAL	P46589 candida alb
7	95	7.1	794	PMSL1_SCHPO	P54280 schizosacch
8	95	7.1	1938	MYHD_HUMAN	O94833 homo sapien
9	93.5	7.0	1024	RIP3_MOUSE	P97434 mus musculus
10	92.5	6.9	459	ZPR1_MOUSE	O62384 mus musculus
11	92.5	6.9	675	NED1_MOUSE	P33215 mus musculus
12	92	6.8	498	MEFA_MOUSE	O60929 mus musculus
13	92	6.8	2774	MAPA_RAT	P34926 rattus norv
14	90.5	6.7	449	AROA_PSEES2	P56952 pseudomonas
15	90	6.7	386	Y364_RICPR	O92087 ticketsia
16	89.5	6.7	823	NSP1_YEAST	P14907 saccharomyc
17	89.5	6.7	885	ASE1_YEAST	P50275 saccharomyc
18	89.5	6.7	2492	ATRX_HUMAN	P46100 homo sapien
19	89	6.6	5020	MEFA_HUMAN	O02078 homo sapien
20	89	6.6	1290	XCPC_XENLA	P50532 xenopus lae
21	89	6.6	2845	APC_MOUSE	O61315 mus musculus
22	88.5	6.6	363	SP43_YEAST	P36094 saccharomyc
23	88.5	6.6	939	ST20_YEAST	O03497 saccharomyc
24	88.5	6.6	1029	RIP3_YEAST	O92087 rattus norv
25	88	6.5	970	Y852_HUMAN	O94639 homo sapien
26	87.5	6.5	284	TPM2_DROME	P09491 drosophila
27	87.5	6.5	558	ORC2_XENLA	O91628 xenopus lae
28	87.5	6.5	642	PHR_NEUCR	P27526 neosporea
29	87.5	6.5	1387	TROP_HUMAN	O12816 homo sapien
30	87.5	6.5	1746	TENA_PIG	O29116 sus scrofa
31	87	6.5	217	EVG1_HUMAN	O94837 homo sapien
32	87	6.5	432	TIG_HAEIN	P40637 haemophilus
33	87	6.5	634	YKCA_CAEEL	P42083 caenorhabdit

ALIGNMENTS

RESULT 1	MEC2_HUMAN	STANDARD:	PRT:	486 AA.
ID	MEC2_HUMAN			
AC	P51608	O15233		
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Methyl-CpG-binding protein 2 (Mecp2 protein) (Mecp2).			
GN	Mecp2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Kudo S., Fukuda M.;			
RL	Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	Tissue=Placenta;			
RL	Thiesen J., Straetling W.H.;			
RL	Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-97130625; PubMed-8976388;			
RT	Vialin A., Aplou F., Vogt N., Dutrillaux B., Malfroy B.;			
RT	"Assignment of the gene for methyl-CpG-binding protein 2 (Mecp2) to human chromosome band Xq28 by in situ hybridization."			
RT	Cytogenet. Cell Genet. 74:293-294(1996).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Reichwald K., Rosenthal A., Kioschis P., Platzer M.;			
RT	"Mapping and sequence analysis of the human Mecp2 locus."			
RT	Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-99299240; PubMed-10369871;			
RT	Coy J.F., Sedlacek Z., Baechner D., Delius H., Poustka A.;			
RT	"A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3'-untranslated region of the methyl-CpG-binding protein 2 gene (Mecp2) suggests a regulatory role in gene expression."			
RT	Hum. Mol. Genet. 8:1253-1262(1999).			
RL	[6]			
RP	SEQUENCE OF 10-486 FROM N.A.			
RC	Tissue=Skeletal muscle;			
RA	MEDLINE-96327611; PubMed-8672133;			
RT	D'Esposito M., Quaderi N.A., Ciccocioppa A., Bruni P., Esposito T.,			
RT	D'Urso M., Brown S.D.M.;			
RT	"Isolation, physical mapping, and Northern analysis of the x-linked human gene encoding methyl CpG-binding protein, Mecp2."			
RT	Mamm. Genome 7:533-535(1996).			
RL	[7]			
RP	SEQUENCE OF 10-486 FROM N.A.			
RA	Reichwald K., Bauer D., Brenner V., Drescher B., Coy J.,			
RA	Kioschis P., Korn B., Nyakatura G., Platzer M., Poustka A.,			

34	87	6.5	861	1	ORC1_HUMAN	O13415 homo sapien
35	86.5	6.4	407	1	YNH8_YEAST	P53939 saccharomyc
36	86.5	6.4	459	1	ZPR1_HUMAN	O75312 homo sapien
37	86.5	6.4	834	1	SRG1_YEAST	O03707 saccharomyc
38	86.5	6.4	986	1	GMI3_RAT	O62839 rattus norv
39	86.5	6.4	1301	1	PTP9_DROME	P35832 drosophila
40	86.5	6.4	2095	1	RPL_MOUSE	P56716 mus musculus
41	86	6.4	310	1	PTPA_PLAUF	P06316 plasmodium
42	86	6.4	411	1	CPRI_PETCR	O99089 petroselinu
43	86	6.4	559	1	PAXI_CHICK	P49024 gallus gall
44	86	6.4	576	1	UNH7_CAEEL	P37806 caenorhabdit
45	86	6.4	622	1	YA41_HUMAN	O9upw0 homo sapien

RA Sandoval N., Rosenthal A.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP VARIANTS RTT TRP-106; CYS-133; SER-155 AND MET-158.
 RX MEDLINE-99438392; PubMed-10508514;
 RA Amir R.E., Van den Veyver I.B., Wan M., Tran C.Q., Francke U.,
 RT Zoghbi H.Y.;
 RT "Rett syndrome is caused by mutations in X-linked MECP2, encoding
 RT methyl-CpG-binding protein 2.";
 RL Nat. Genet. 23:185-188(1999).
 RN [9]
 RP VARIANT RTT VAL-140.
 RX MEDLINE-20465115; PubMed-11007980;
 RA Orrico A., Lam C., Galli L., Dotti M.T., Hayek G., Tong S.F.,
 RT Poon P.M., Zappella M., Federico A., Sorrentino V.;
 RT "MECP2 mutation in male patients with non-specific X-linked mental
 RT retardation.";
 RL FRS Lett. 481:285-288(2000).
 RN [10]
 RP VARIANTS RTT W-106; F-124; C-13; C-134; R-152; M-158 AND C-306.
 RX MEDLINE-20439334; PubMed-10991688;
 RA Ohta K., Matsushita T., Yamashita Y., Fukuda T., Kuwajima K.,
 RT Horinouchi I., Nagamitsu S., Iwanaga R., Kimura A., Omori I., Endo S.,
 RA Mori K., Kondo I.;
 RT "Mutation analysis of the methyl-CpG binding protein 2 gene (MECP2) in
 RT patients with Rett syndrome.";
 RL J. Med. Genet. 37:608-610(2000).
 RN [11]
 RP VARIANTS RTT R-101; W-106; M-158 AND C-306, AND VARIANT K-397.
 RX MEDLINE-20439335; PubMed-10991689;
 RA Hampson K., Woods C.G., Latif F., Webb T.;
 RT "Mutations in the MECP2 gene in a cohort of girls with Rett
 RT syndrome.";
 RL J. Med. Genet. 37:610-612(2000).
 RN [12]
 RP VARIANT RTT VAL-140.
 RX MEDLINE-21664240; PubMed-11805248;
 RA Dotti M.T., Orrico A., De Stefano N., Battisti C., Sicurelli F.,
 RT Severi S., Lam C.W., Galli L., Sorrentino V., Federico A.;
 RT "A Rett syndrome MECP2 mutation that causes mental retardation in
 RT men.";
 RL Neurology 58:226-230(2002).
 CC -1- FUNCTION: CHROMOSOMAL PROTEIN THAT BINDS TO METHYLATED DNA. IT CAN
 CC BIND SPECIFICALLY TO A SINGLE METHYL-CpG PAIR. IT IS NOT
 CC INFLUENCED BY SEQUENCES FLANKING THE METHYL-CpG. MEDIATES
 CC TRANSCRIPTIONAL REPRESSION THROUGH INTERACTION WITH HISTONE
 CC DEACETYLASE AND THE COREPRESSOR SIN3A.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. COLOCALIZED WITH METHYL-CpG IN THE
 CC GENOME.
 CC -1- TISSUE SPECIFICITY: PRESENT IN ALL ADULT SOMATIC TISSUES TESTED.
 CC -1- DISEASE: DEFECTS IN MECP2 ARE THE CAUSE OF RETT SYNDROME (RTT), AN
 CC X-LINKED DOMINANT DISEASE. RTT IS A PROGRESSIVE NEUROLOGIC
 CC DEVELOPMENTAL DISORDER AND ONE OF THE MOST COMMON CAUSES OF MENTAL
 CC RETARDATION IN FEMALES. PATIENTS APPEAR TO DEVELOP NORMALLY UNTIL
 CC 6 TO 18 MONTHS OF AGE, THEN GRADUALLY LOOSE SPEECH AND PURPOSEFUL
 CC HAND MOVEMENTS AND DEVELOP MICROCEPHALY, SEIZURES, AUTISM, ATAXIA,
 CC INTERMITTENT HYPERVENTILATION, AND STEREOTYPIC HAND MOVEMENTS.
 CC AFTER INITIAL REGRESSION, THE CONDITION STABILIZES AND PATIENTS
 CC USUALLY SURVIVE INTO ADULTHOOD.
 CC -1- SIMILARITY: CONTAINS 1 A.T HOOK DNA-BINDING REPEAT.
 CC -1- SIMILARITY: CONTAINS 1 METHYL-BINDING DOMAIN (MBD).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L37298; AAC32737.1; -
 CC DR EMBL: Y12643; CAA73190.1; -
 CC DR EMBL: X99686; CAA68001.1; -
 CC

DR EMBL: AF030876; AAC08757.1; -
 DR EMBL: AF031078; AAC08758.1; -
 DR EMBL: AJ132917; CAB6446.1; -
 DR EMBL: X89430; CAA61599.1; -
 DR EMBL: X94628; CAA64331.1; -
 DR TRANSFAC: T04936; -
 DR Genew: HGNC:6990; MECP2.
 DR MIM: 300005; -
 DR MIM: 312750; -
 DR InterPro: IPR000637; AT_hook.
 DR InterPro: IPR001739; Methyl-CpG_bind.
 DR Pfam: PF01429; MBD; 1.
 DR SMART: SM00384; AT_hook; 1.
 DR SMART: SM00391; MBD; 1.
 KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;
 KW Disease mutation; Polymorphism.
 KM DOMAIN 96 149
 FT DOMAIN 277 283
 FT DOMAIN 366 372
 FT DOMAIN 384 393
 FT VARIANT 101 101
 FT VARIANT 106 106
 FT VARIANT 124 124
 FT VARIANT 133 133
 FT VARIANT 134 134
 FT VARIANT 140 140
 FT VARIANT 152 152
 FT VARIANT 155 155
 FT VARIANT 158 158
 FT VARIANT 201 201
 FT VARIANT 306 306
 FT VARIANT 397 397
 FT CONFLICT 72 75
 FT CONFLICT 290 290
 SQ SEQUENCE 486 AA; 52440 MW; EB6A33233AEDA566 CRC64;
 Query Match 13.2%; Score 177; DB 1; Length 486;
 Best Local Similarity 28.4%; Pred. No. 6; 7e-06;
 Matches 75; Conservative 37; Mismatches 86; Indels 66; Gaps 14;
 QY 2 DCPALPPGKMKKEVIRKSGLSAKKSDVYFSPGKFRFSKPOLARY---LGNT-VDLSSF 57
 DB 96 DDFTLPGWGRKRLKORSGRSAGKYVYLLNPGKAFRSKVELLAFKEVVDGTSIDPNDF 155
 QY 58 DFR-TGKMPKSKLOKKORLRLNDP---LNONKGR-DINTPLIPROTASIFKOPYTKYT 111
 DB 156 DFTYTGGSRSKREQRPKPKSPKAPGCRGGRKSGGTTTPKATSEGVQ--VKRVL 213
 QY 112 NHPKSKVSDPQRNEDPQLFEWKRLQGISADVTE-QITK-----TMELPKG 159
 DB 214 E-----KSPGKLLVKKMFPQTSFGKAEAGGATTSQVWVKKPKRRAADPPAIPK- 266
 QY 160 LQGVGPGSSNDETLISAVASLHSHSSAPITQGVANAEKNPAVWLNSTQPLCKAFIVTDSD 219
 DB 267 KRGRKPG-----SVVAAA-----AAEAKKKA-----YKESS 292
 QY 220 IRKQEEVQVQRK-KLEELAMD 242
 DB 293 IRSVQETVLPFKKKRTREIVSIEV 316

RESULT 2
MEC2_MOUSE STANDARD: PRT: 484 AA.
AC 0922D6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methyl-Cpg-binding protein 2 (Mecp2-2 protein) (Mecp2).
GN MECP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=98449942; PubMed=9774669;
RA Hendrich B., Bird A.;
RT "Identification and characterization of a family of mammalian methyl-
Cpg binding proteins."
RT Mol. Cell. Biol. 18:6538-6547(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99299240; PubMed=10369871;
RA Coy J.F., Sedlacek Z., Baechner D., Delius H., Pousetka A.;
RT "A complex pattern of evolutionary conservation and alternative
polyadenylation within the long 3'-untranslated region of the methyl-
Cpg-binding protein 2 gene (Mecp2) suggests a regulatory role in gene
expression."
RT Hum. Mol. Genet. 8:1253-1262(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Reichwald K., Thiesse J., Wlehe T., Kioschis P., Straetling W.H.,
RA Rosenthal A., Platter M.;
RT "Comparative analysis of the methyl Cpg binding protein 2 locus in man
and mouse reveals new untranslated sequences."
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CHROMOSOMAL PROTEIN THAT BINDS TO METHYLATED DNA. IT CAN
CC BIND SPECIFICALLY TO A SINGLE METHYL-CPG PAIR. IT IS NOT
CC INFLUENCED BY SEQUENCES FLANKING THE METHYL-CPGs. MEDIATES
CC TRANSCRIPTIONAL REPRESSION THROUGH INTERACTION WITH HISTONE
CC DEACETYLASE AND THE COMPRESSOR SIN3A (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR. COLOCALIZED WITH METHYL-CPG IN THE
CC GENOME.
CC -1- SIMILARITY: CONTAINS 1 A.T HOOK DNA-BINDING REPEAT.
CC -1- SIMILARITY: CONTAINS 1 METHYL-BINDING DOMAIN (MBD).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF072251; AAC68880.1; -
DR EMBL: AJ132922; CAB46495.1; -
DR EMBL: AF121351; AAF23116.1; -
DR EMBL: AF158181; AAF33024.1; -
DR MCD: MGI:99918; Mecp2.
DR InterPro: IPR000637; AT_hook.
DR InterPro: IPR001739; Methyl-Cpg_bind.
DR Pfam: PF01429; MBD.1.
DR PRINTS: PR00929; ATHOOK.
DR SMART: SM00391; MBD.1.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DOMAIN 56 149 MBD.
FT DOMAIN 277 283 POLY-ALA.
FT DOMAIN 366 372 POLY-HIS.
FT DOMAIN 384 391 POLY-PRO.
FT DOMAIN 440 443 POLY-THR.
SO SEQUENCE 484 AA; 52307 MW; 62FD228F0118A49F CRC64;

Query Match 12.4%; Score 167; DB 1; Length 484;
Best Local Similarity 26.3%; Pred. No. 3.6e-05;
Matches 71; Conservative 31; Mismatches 90; Indels 78; Gaps 12;
OY 2 DCPALPQGMKKEVIRKSGLSKSDVYFSPSGKRRKPOLARY---LGNT-VDLSSF 57
DB 96 DPTTLPEGWTRKLRKQKSGRSKGYDYLINPOKAFRSKVELLAVFEKVGDSLDPNDF 155
OY 58 DFR-TGKMPSKLOKKNQ-----RLRNDPLNQNKGKPDLTTLPIRQTASIFK 104
DB 156 DFTVYGRSGSPSRREQPPKPKSPKAPGTGGRPRPGSGTGKPKAASGVQKRYLEK 215
OY 105 QPVTKYVTHNPSNKKYKSDPORNNEQPROLFEWKRLOGLSASDVTEQITK-----T 153
DB 216 SPGKLVRKMP---FQASPGGKGE-----GGGATTSAGQVMVVIKRRGRKKAADP 261
OY 154 MELPKLOGVPGSNDFTLLSAVASALHTSSAPITGQVSAVAEKNPVWMLTSPICKAF 213
DB 262 QAIKPK-KGGRKPG-----SVAAA-----AAEAKKA----- 287
OY 214 IVTDEIRKQGEERYQOVRK-KLEBALMADI 242
DB 288 -VKESIRSHVETVLPIKRRKRTVSIEV 316
RESULT 3
MEC2_RAT STANDARD: PRT: 492 AA.
AC Q00566;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methyl-Cpg-binding protein 2 (Mecp2-2 protein) (Mecp2).
GN MECP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92298389; PubMed=1606614;
RA Lewis J.D., Meehan R.R., Henzel W.J., Maurer-Fogy I., Jeppesen P.,
RA Klein F., Bird A.;
RT "Purification, sequence, and cellular localization of a novel
RT chromosomal protein that binds to methylated DNA."
RT Cell 69:905-914(1992).
CC -1- FUNCTION: CHROMOSOMAL PROTEIN THAT BINDS TO METHYLATED DNA. IT CAN
CC BIND SPECIFICALLY TO A SINGLE METHYL-CPG PAIR. IT IS NOT
CC INFLUENCED BY SEQUENCES FLANKING THE METHYL-CPGs. MEDIATES
CC TRANSCRIPTIONAL REPRESSION THROUGH INTERACTION WITH HISTONE
CC DEACETYLASE AND THE COMPRESSOR SIN3A (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR. COLOCALIZED WITH METHYL-CPG IN THE
CC GENOME.
CC -1- TISSUE SPECIFICITY: PRESENT IN ALL ADULT SOMATIC TISSUES TESTED.
CC -1- SIMILARITY: CONTAINS 1 A.T HOOK DNA-BINDING REPEAT.
CC -1- SIMILARITY: CONTAINS 1 METHYL-BINDING DOMAIN (MBD).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M94064; AAA41584.1; -
DR FIR: A41907; A41907.
DR InterPro: IPR001739; Methyl-Cpg_bind.
DR Pfam: PF01429; MBD.1.
DR SMART: SM00391; MBD.1.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DOMAIN 96 149 MBD.

FT DOMAIN 366 372 POLY-HIS.
 FT DOMAIN 384 391 POLY-PRO.
 FT DOMAIN 443 451 POLY-THR.
 SQ SEQUENCE 492 AA; 53047 MW; A67E05068BA2D38 CRC64;
 Query Match 12.4%; Score 167; DB 1; Length 492;
 Best Local Similarity 28.1%; Pred. No. 3,7e-05;
 Matches 74; Conservative 37; Mismatches 88; Indels 64; Gaps 13;
 QY 2 DCPALPCKKKKEEVIRKSLGKSDVYEFSPSGKPKRKKPOLARY--LGNT-VDLSSE 57
 DB 96 DDEPLPEGWTRKIKOKSSRSAGKTDVLLINQKAFPSKVELLAYPEKVDTSIDNDNF 155
 QY 58 DFR-TGKMPKSKLQKNKORLRNDP---LNQKGRPDINTLPIRQNASIFKQPVYKVTN 112
 DB 156 DFLVYGRGSPSRKQPKPKSPKAPGGRGRGRKPSGGRPKRAAASEGVQ--VKRYLE 214
 QY 113 HPSNKKVSDPQRMNEQPRLEFKRLQGISADVTE-OLIK-----TWELPKGL 160
 DB 215 -----KSPGKILVKMPFQASPGGEGGATTSQAVYIKRPGKRRRAADPOAIRK-K 267
 QY 161 QGVGSGNDETLLSAVASALHTSSAPITGVSAVEKNPAVWLNTSOLCKAFIYTDEDI 220
 DB 268 RGRKPG-----SVYAAA-----AAEAKKKA-----YKESST 293
 QY 221 RKOERYVOYKR-KLEALMADI 242
 DB 294 RSVQETVLPKIKRKRTRVTSIEV 316
 RESULT 4
 MS1_DROME STANDARD; PRT; 1039 AA.
 ID MS1_DROME STANDARD; PRT; 1039 AA.
 AC P50535;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Male-specific lethal-1 protein.
 GN MS1-1.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE OF 85-1039 FROM N.A.
 RP STRAIN=Canton-S;
 RX MEDLINE=93314941; PubMed=8325488;
 RA Palmer M.J., Mergner V.A., Richman R., Manning J.E., Kuroda M.I.,
 RA Lucchesi J.C.;
 RA "The male-specific lethal-one (msl-1) gene of Drosophila melanogaster
 RT encodes a novel protein that associates with the X chromosome in
 RT males.";
 RL Genetics 134:545-557(1993).
 RN [2]
 RP REVISIONS, SEQUENCE FROM N.A.
 RX MEDLINE=95300219; PubMed=781064;
 RA Kelley R.L., Solovayeva I., Lyman L.M., Richman R., Solovay V.,
 RA Kuroda M.I.;
 RA "Expression of msl-2 causes assembly of dosage compensation
 RT regulators on the X chromosomes and female lethality in Drosophila.";
 RL Cell 81:867-877(1995).
 CC -1- FUNCTION: THE MSL PROTEINS ARE ESSENTIAL FOR ELEVATING
 CC TRANSCRIPTION OF THE SINGLE X CHROMOSOME IN THE MALE (X CHROMOSOME
 CC DOSAGE COMPENSATION). MSL-1 IS A PIONEER PROTEIN. MLE, MSL-1 AND
 CC MSL-3 ARE CO-LOCALIZED ON THE X CHROMOSOME. EACH OF THE MSL
 CC PROTEINS REQUIRES ALL THE OTHER MSLS FOR WILD-TYPE X-CHROMOSOME
 CC BINDING.
 CC -1- SUBUNIT: MSL-1 SEEMS TO FORM A TIGHT COMPLEX WITH MSL-2.
 CC -1- SUPRACELLULAR LOCATION: NUCLEAR; MSL-1 IS ASSOCIATED WITH HUNDREDS
 CC OF DISCRETE SITES ALONG THE LENGTH OF THE X CHROMOSOME IN MALES
 CC AND NOT IN FEMALES, AND IS ALSO ASSOCIATED WITH 10-20 AUTOSOMAL
 CC SITES IN MALES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: L42514; AAA98918.1; -;
 DR Flybase: FBgn0005617; msl-1.
 KW Nuclear protein.
 FT CONFLICT 188 193 PLPPAA -> HCHLLP (IN REF. 1).
 FT CONFLICT 492 492 L -> S (IN REF. 1).
 FT CONFLICT 670 670 I -> M (IN REF. 1).
 SQ SEQUENCE 1039 AA; 117412 MW; 4759E95EE6E9F14 CRC64;
 Query Match 8.0%; Score 107.5; DB 1; Length 1039;
 Best Local Similarity 23.6%; Pred. No. 2.4;
 Matches 48; Conservative 26; Mismatches 58; Indels 71; Gaps 7;
 QY 73 KORLRNDPLNOKGR-----DLNTLPIRQNASIFKQPVYKV-TNPSNKKVKS 120
 DB 558 KETLRKQPDAPKHLPRKAVAPRYVYKTSRSTFLPRANTADIKDAPQKVIANNHSTQRT 617
 QY 121 DP---ORANQOPRO--LFWEKRLQGISADVTEOLIKTMELPKGLQVPGSNDT--- 171
 DB 618 DPKYQRLQYKIRQYEMHMDMTGSSAPSDIRKQ-----KNVDPVSTPEYKTIK 666
 QY 172 ----LISAVASALHTSSAPITGVSAVEKNPAVWLNTSOLCKAFIYTDIDKOEERY 227
 DB 667 SKSILVNDKKTSETSQS-----DQELDV 691
 QY 228 QGVKRLKEALMADILSRADTE 250
 DB 692 ETVRKRLEHLKKEILLSHSSQ 714
 RESULT 5
 YAP1_CHICK STANDARD; PRT; 448 AA.
 ID YAP1_CHICK STANDARD; PRT; 448 AA.
 AC P46936;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 65 kDa Yes-associated protein (YAP65).
 GN YAP1 OR YAP65.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCBI_TaxID=9031;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=White Leghorn;
 RX MEDLINE=94309887; PubMed=8035999;
 RA Sudol M.;
 RA "Yes-associated protein (YAP65) is a proline-rich phosphoprotein that
 RT binds to the SH3 domain of the Yes proto-oncogene product.";
 RL Oncogene 9:2145-2152(1994).
 CC -1- FUNCTION: BINDS TO THE SH3 DOMAIN OF THE YES KINASE. ALSO BINDS TO
 CC OTHER SIGNALING MOLECULES THAT CONTAIN SH3 DOMAINS INCLUDING NCK,
 CC CRK AND SRC.
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES.
 CC -1- SIMILARITY: CONTAINS 1 OR 2 WW DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL: X76483; CAA54021.1;
 DR InterPro: IPR005153; Mbth.
 DR InterPro: IPR002349; MW.
 DR InterPro: IPR001202; MW_Rsp5_WMP.
 DR Pfam: PF00397; MW; 1.
 DR Pfam: PF03621; Mbth; 1.
 DR PRINTS: PR00403; WMDOMAIN.
 DR SMART: SM00456; MW; 1.
 DR PROSITE: PS01159; MW_DOMAIN_1; 1.
 DR PROSITE: PS00020; MW_DOMAIN_2; 1.
 DR Phosphorylation.
 DR DOMAIN 169 202 MW.
 FT SEQUENCE 448 AA; 47822 MW; 719CC8D0F879A3BD CRC64;
 SQ

Query Match
 Best Local Similarity 7.2%; Score 97; DB 1; Length 448;
 Matches 64; Conservative 35; Mismatches 114; Indels 78; Gaps 10;

QY 2 DCPALPGKKKEEVIRKSGLSAGKSDVYFSPSGKFKRSPQALRYLGNTVDSLSPFRT 61
 DB 168 DVP-LPPGWEAK-----TPSG-----RYFLNHIDOTTWDDP 200
 QY 62 GKMPSTKLOKKNQRLRNPLNONGKRPDLNTLPIRQTASTFKQPVTKVTHPSNKV--- 118
 DB 201 RKAMS--QMTVAPTSPVQNNLMSASAMNORISQAPV-KQPLAPQSPQGVWG 257
 QY 119 KSDPQRMNEQPRQLEFMEKRLQGLSASDVTEQ-----IKTMELPKLQG--VGPG--- 166
 DB 238 SSSNQOQMRLOQLQMEKERLRKHQELRLQRLSOLPTEMDGGSSQVPSGKSGQ 317
 QY 167 -----SNDLTLASVASALHTSSAPITGQVSAVEKNPAVWLNTSOPCKAFITDED 219
 DB 318 LRTMTNNSDPLNSGYHSRDESTDGSLMSYSYVPTPDDFLNSVDMDTDDTSISQSV 377
 QY 220 IRKQERY-----QOVRKLEALMDILS 244
 DB 378 IPSHQNRPPDYLEAPGNVNDGLTEGDKNIEBELMPSLQELSLSDILN 428

RESULT 6

ADFL_CANAL STANDARD; PRT; 612 AA.
 AC P46589;
 DT 01-OCT-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Adherence factor (Adhesion and aggregation mediating surface antigen).
 GN ADFL OR AAF1.
 OS Candida albicans (yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 ON NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE OF 35-612 FROM N.A.
 RC STRAIN=ATCC 36082;
 RA Edwards J.E.;
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Jiang W., Finkler A., Koltin Y.;
 RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: SURFACE ANTIGEN MEDIATING ADHESION AND AGGREGATION IN
 CC S.CEREVISIAE.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: U18983; AAA62506.1; -
 DR EMBL: U44747; AAA6758.1; -
 KW Cell adhesion.
 FT DOMAIN 64 69 POLY-GLN.
 FT DOMAIN 100 103 POLY-SER.
 FT DOMAIN 117 134 POLY-GLN.
 FT DOMAIN 269 272 POLY-ASN.
 FT DOMAIN 517 520 POLY-PRO.
 FT DOMAIN 579 582 POLY-PRO.
 FT DOMAIN 596 605 POLY-GLN.
 FT CONFLICT 142 142 F -> N (IN REF. 1).
 FT CONFLICT 451 451 T -> I (IN REF. 1).
 SQ SEQUENCE 612 AA; 68794 MW; 77B60B45C35C1B23 CRC64;

Query Match
 Best Local Similarity 7.1%; Score 95; DB 1; Length 612;
 Matches 51; Conservative 38; Mismatches 91; Indels 48; Gaps 10;

QY 36 KFRSKPOLARYLGNTVDSLSPFRTGKMPSTKLOKKNQRLRNPLNONGKRPDLNTLP 95
 DB 131 QOMQOPQNMQEFNDNTI--PNVLMNQITSPQOTTTQ--PNISYNYSTQPOLQSAQ 166
 QY 96 IQGTASIRKQPVTKVTHPSNKVSDPQRMNEQPRQLEFMEKRLQGLSASDVTEQIIKTM 155
 DB 187 ISHSQ---PQPPQATQPSRNSRS--QTSFKPR--GSHQVSGSGSTGAK---KQSA 235
 QY 156 LFKGLQVGPSSNDLTLASVASALHTSSAPITG-----QVSAVEKNPAVWLNTSOPLC 210
 DB 236 ITSGSTGTGPARNMDTGTFTSVANSTSTTTMTNNKLSVASPVN---VIVANL----- 287
 QY 211 KAFITDEDIRKQERYQOVRKLEALMDILSRAADTEEMDIEMDS 258
 DB 288 -----PERLQOV-----LPAPLSRAVPRDVTYVNLIS 315

RESULT 7

PMS1_SCHPO STANDARD; PRT; 794 AA.
 ID PMS1_SCHPO
 AC P54280;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA mismatch repair protein pms1.
 GN PMS1 OR SPAC19G12.02C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetales;
 OC Schizosaccharomycetes.
 ON NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA MEDLINE=97403304; PubMed=9258673;
 RA Schar P., Baur M., Schneider C., Kohl J.;
 RT "Mismatch repair in Schizosaccharomyces pombe requires the mult
 RL homologous gene pms1: molecular cloning and functional analysis.";
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltrielli T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Meljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambolt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Potashin J.,
 RA Cernutti L., Lowe T., McCombie W.R., Paulsen I., Potashin J.,
 RA Shpitskiy G.V., Ussery D., Barrett B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES
 CC IN DNA.
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MTHL/HEX FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: X96581; CAB5400.1; -
 CC EMBL: Z97209; CAB10113.1; -
 CC HSSP: P23367; IBKN.
 CC InterPro: IPR003594; ATPbind_ATPase.
 CC InterPro: IPR002099; DNA_mis_repair.
 CC Pfam: PF01119; DNA_mis_repair; 1.
 CC Pfam: PF02518; HATase_c; 1.
 CC TIGRfam: TIGR00585; multi; 1.
 CC PROSITE: PS00058; DNA_MISMATCH_REPAIR_1; 1.
 CC DNA repair.
 CC
 CC SEQUENCE 794 AA; 88009 MW; ASD46FFFA077D8DC CRC64;
 SQ
 Query Match 7.1%; Score 95; DB 1; Length 794;
 Best Local Similarity 23.7%; Pred. No. 14;
 Matches 46; Conservative 32; Mismatches 62; Indels 54; Gaps 9;
 QY 18 KSGLSAGKSDVYFSPSGKFE---RSKPQIARYLGNVTDLSSDFPTGKMMPSKLOKNO 74
 DB 373 ESERSDDSSFSYKRSKPKRLVETAPAISTVAEGASIA---GVSKPLPERLOKDSM 428
 QY 75 RLRLDPLNQN-----KGGPDL---NTLPIRQOTASI---FKQVTKVTHPSNKKVKS 120
 DB 429 R-RSSPLNEKYTASSEPMKKKLLAFASSTSTSMOKITDSSFPLKQPIKNPSSNNLLN 487
 QY 121 DPGRMNEQPROLFEWKRLOGLSASDVTEQIKTWELPKGLQGVGPSNDETLISAVASAL 180
 DB 488 DP-----SPASTPVAKTINLINE-IESVHNAESVSTL----- 517
 QY 181 HTSSAPITGOVSAA 194
 DB 518 --SSTPRTQTSVA 529
 RESUME 8
 MYHD_HUMAN STANDARD; PRT; 1938 AA.
 AC Q9UKX3; G95252;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, skeletal muscle, extraocular (MyHC-ec).
 GN MYH3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE-Extraocular muscle;
 RX MEDLINE=99318869; PubMed=10386558;
 RA Weiss A., Schaffino S., Leinwand L.A.;
 RT "Comparative sequence analysis of the complete human sarcomeric myosin
 RT heavy chain family: implications for functional diversity."
 RL J. Mol. Biol. 290:61-75(1999).
 RN [2]
 RP SEQUENCE OF 1917-1938 FROM N.A.
 RP TISSUE-Extraocular muscle;
 RX MEDLINE=99026150; PubMed=9806854;
 RA Winters L.M., Briggs M.M., Schachar F.;
 RT "The human extraocular muscle myosin heavy chain gene (MYH13) maps to
 RT the cluster of fast and developmental myosin genes on chromosome 17."
 RL Genomics 54:188-189(1998).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MYC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMY) AND 1 HEAVY MEROMYOSIN (HMY). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: AF111782; AAD29948.1; -
 CC EMBL: AF075248; AAC83241.1; -
 CC HSSP: P13538; 2MYS.
 CC GeneW: HGNC:7571; MYH13.
 CC MTR: 603487; -
 CC InterPro: IPR000048; IQ_region.
 CC InterPro: IPR004009; Myosin_N.
 CC InterPro: IPR002928; Myosin_tail.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF00063; myosin_head; 1.
 CC Pfam: PF00612; IQ; 2.
 CC Pfam: PF01576; Myosin_tail; 1.
 CC Pfam: PF02736; Myosin_N; 1.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC PRODOM: PD000355; myosin_head; 1.
 CC SMART: SM00015; IQ; 1.
 CC SMART: SM00242; MYSC; 1.
 CC PROSITE: PS50096; IQ; 1.
 KW Myosin: Muscle protein; coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW MultiGene family.
 FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
 FT DOMAIN 785 814 IQ.
 FT DOMAIN 843 1938 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 761 775 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 130 130 METHYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 699 699 ALKYLATION (SH-2) (POTENTIAL).
 FT MOD_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).
 FT MOD_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).
 FT SEQUENCE 1938 AA; 223678 MW; 1F6D006416381C05 CRC64;
 SQ
 Query Match 7.1%; Score 95; DB 1; Length 1938;

Best Local Similarity 20.5%; Pred. No. 44;
Matches 61; Conservative 49; Mismatches 139; Indels 40; Gaps 7;

OY 2 DCALPPGKKKEEYIR---KSGLSAGSDV-----YFSPSGKKFRSKPOLARYLGNVVD 53
Db 1343 DCCLRLROYEEDEAKLORALSKANSEVAMQKTYETDAIORTEELEAKKLAORLO 1402
OY 54 LSSFDFTGKMMPSKLOKNKORLND-----PLNQNGKRDLM 92
Db 1403 EAEKFTANSKCSASLEKTKORLOGEVEDLMDRLERSHTACATLDKKORFDDVLAEMKQ 1462
OY 93 TLPDRTASIFKOPVTKVTNHPNSKVKSDPQRNEDQPROLFMEKRLQGLSADVTEQIIR 152
Db 1463 KLEESQAELEAAQKESRSLSTELFKRNNAYEEVVDQLETLRNKNMLQGEISDLTEQIANE 1522
OY 153 TMLPPLGLOGVPGSNDETLLSAVALHTSSAPITGOVSAVEKNPAYWLNTSQ---PL 209
Db 1523 T---GKNLQ---EAEKTKLVEQEKSDLOVALVEEGSLHEBSKILRVQLELSQVKSBL 1576
OY 210 GKAFIYDDEDIRKOEERVOQVRKKLEALMADILSR-----AADTEMDIEM 256
Db 1577 DRVIEKDEIEQLKNSQRAALALOSVDIAEIRSRNDALRLKKMEGLNEMETOL 1633

RESULT 9

ID RIP3_MOUSE STANDARD: PRT: 1024 AA.
AC P97434:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Rho-interacting protein 3 (p16Rip) (RIP3).
GN RHOFIP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97344280; PubMed=9199174;
RA Gebbink M.F.B.G., Kranenburg O., Poland M., van Horck F.P.G.,
RT Housa B., Moolenaar W.H.;
RT Identification of a novel, putative Rho-specific GDP/GTP exchange
RT factor and a Rho-binding protein: control of neuronal morphology.;
RT J. Cell Biol. 137:1603-1613(1997).
CC -!- FUNCTION: RHO-BINDING PROTEIN INVOLVED IN CONTROL OF THE ACTIN
CC CYTOSKELETON. OVEREXPRESSION PROMOTES NEURONAL CELL FLATTENING AND
CC NEURITE OUTGROWTH, PROBABLY BY COUNTERACTING RHOA-MEDIATED
CC SIGNALING.
CC -!- TISSUE SPECIFICITY: HIGHLY ENRICHED IN THE BRAIN.
CC -!- SIMILARITY: CONTAINS 2 PH DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: U73200; ABI18198.1; -
CC DR MGI: MGI:1349438; Rho1d3.
CC DR InterPro: IPR001849; PH.
CC DR Pfam: PF00169; PH: 2.
CC DR SMART: SM00233; PH: 2.
CC DR PROSITE: PS00003; PH_DOMAIN: 2.
CC KM Guanine-nucleotide releasing factor; Repeat; Coiled coil.
CC FT DOMAIN 43 150
CC FT DOMAIN 386 482
CC FT DOMAIN 164 172
CC FT DOMAIN 185 190
CC FT DOMAIN 285 296
CC PRO-RICH.
CC POLY-RICH.
CC PRO-RICH.

FT DOMAIN 671 976 COILED COIL (POTENTIAL).
SQ SEQUENCE 1024 AA, 116363 MW, A91BDB955E45E02 CRC64;

Query Match 7.0%; Score 93.5; DB 1; Length 1024;
Best Local Similarity 19.5%; Pred. No. 25;
Matches 61; Conservative 53; Mismatches 132; Indels 67; Gaps 11;

OY 5 ALPPGKKKEEYIRKSGLSAGSDVYFSPSGKKFRSKPOLARYLGNVVDLSFPDR--- 60
Db 492 SLPEGKKNSTSPETCSRSTKQEAEPGCEPDEQKKSRARRRRREGSKSTFWMAEPRPQQ 551
OY 61 -----KGKAPDLNTT 93
Db 552 ALAQSRASVSSDSGDCCEAEPEGLEREPARRRPFGLDTIDSGMEDVALR 611
OY 94 LPIRQTASIFKOPVTKVTN-HPS-----NKYSDPQRNEDQPROLFMEKRLQGLSA 143
Db 612 MDIDRSPGLGTPDLKTQNVHVEIQRWQYETTPLEKQVPLAHLSLSDSERLST 671
OY 144 SVTQIITKTMELPKGLGVGPGSN---DETLLSAVALHTSSAPITGOV-SAVEKN 198
Db 672 HELTSLERKELE-----OSQKEASDLLEQNRLDOLRVALGRQSAREGVVLQATCERG 726
OY 199 PAVWLNTSOPICKAFIYDDEDIRKOEERVOQVRKKLEALMADILS-----RAADT 249
Db 727 FAAMETHQKKI-----EDLQROHORELEKLEEKDRLLAEETVAITSAIEAMKNHR 779
OY 250 EEMDIEMDSGDEA 262
Db 780 EEMERELEKSQRS 792

RESULT 10

ID ZPRL_MOUSE STANDARD: PRT: 459 AA.
AC Q62384:
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc-finger protein ZpR1 (Zinc finger protein 259).
GN ZNP259 OR ZFP259 OR ZPRL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Galcheva-Gargova Z., Konstantinov K.N., Wu I.-H., Klier F.G.,
RT Barrett T., Davis R.J.;
RT "Binding of zinc finger protein ZpR1 to the epidermal growth
RT factor receptor.";
RT Science 272:1797-1802(1996).
CC -!- SUBCELLULAR LOCATION.
CC MEDLINE=98437195; PubMed=9763455;
CC Galcheva-Gargova Z., Gangwan L., Konstantinov K.N., Mikrut M.,
CC Theroux S.J., Enoch T., Davis R.J.;
CC "The cytoplasmic zinc finger protein ZpR1 accumulates in the nucleolus
CC of proliferating cells.";
CC Mol. Biol. Cell 9:2963-2971(1998).
CC -!- FUNCTION: MAY BE A SIGNALING MOLECULE THAT COMMUNICATES MITOGENIC
CC SIGNALS FROM THE CYTOPLASM TO THE NUCLEUS.
CC -!- SUBUNIT: BINDS TO THE EGF AND PDGF RECEPTORS. BINDS TO THE
CC ELONGATION FACTOR 1-ALPHA.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC; TRANSLOCATE TO THE NUCLEUS
CC AFTER TREATMENT WITH MITOGENS.
CC -!- SIMILARITY: BELONGS TO THE ZPRL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U30823; AAA74030.1; -
 DR HSPSP: P11831; 1SR5.
 DR TRANSFAC: T00505; -
 DR MGD: MGI:99532; Mef2a.
 DR InterPro: IPR002100; TF_MADSbox.
 DR Pfam: PF00319; SRP-TP; 1.
 DR PRINTS: PR00404; MADSDOMAIN.
 DR SMART: SM00432; MADS; 1.
 DR PROSITE: PS00350; MADS_BOX_1; 1.
 DR PROSITE: PS00066; MADS_BOX_2; 1.
 KW Transcription regulation; Nuclear protein; DNA-binding; Activator;
 KM Multigene family.
 FT DOMAIN 3 57 MADS.
 FT DNA_BIND 58 86 MEF2-TYPE (POTENTIAL).
 FT DOMAIN 254 257 POLY-PRO.
 FT DOMAIN 288 293 POLY-GLU.
 FT DOMAIN 419 423 POLY-HIS.
 FT DOMAIN 448 455 POLY-SER.
 SQ SEQUENCE 498 AA; 53724 MW; 590678D1BD1B3723 CRC64;

Query Match 6.8%; Score 92; DB 1; Length 498;
 Best Local Similarity 27.4%; Pred. No. 13;
 Matches 51; Conservative 18; Mismatches 51; Indels 66; Gaps 11;

QY 32 SPSPGKFF---RSKPOLAVYLGNTYDLSFDRTKMMPSKLOKRLNDP-----LN 82
 DB 221 SPVGNFVNSRASPWL---IGNTGANS---LQKVMPTK-----SPPPGGSG 263
 QY 83 QNKGKPDINTLPIROTASTIFKQPVTKYTNHPSKVKSDPQRNNEQPRQLFWKRIQGLS 142
 DB 264 MNSRKPDLRLVAIP-----PSKGMPLSEEE---LELNQRIS 300
 QY 143 ASDYTEQI---IKTMELPKGLQV-----GPSNDETLTLLSAVASALHTSSAP---I 187
 DB 301 SSOATOPATPVPVSVTPPSLP--QGLVSGAMPTAVNTDYLTSADLSAQGTSPGMLS 358
 QY 188 TGOVSA 193
 DB 359 LGQASA 364

RESULT 13

MAPA_RAT STANDARD; PRT: 2774 AA.

AC P34936;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Microtubule-associated protein 1A (MAP 1A) [Contains: MAP1 light chain
 GN LC2].
 GN MAP1A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92355629; PubMed=1379599;
 RA Langkopf A., Hammarback J.A., Mueller R., Vallee R.B., Garner C.C.;
 RT "Microtubule-associated proteins 1A and LC2. Two proteins encoded in
 J. Biol. Chem. 267:16561-16566(1992).
 CC -1- FUNCTION: Structural protein involved in the filamentous cross-
 CC bridging between microtubules and other skeletal elements.
 CC -1- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate

CC with MAP1A and MAP1B proteins.
 CC -1- TISSUE SPECIFICITY: BRAIN, HEART AND MUSCLE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED LATE DURING NEURONAL DEVELOPMENT
 CC APPEARING WHEN AXONS AND DENDRITES BEGIN TO SOLIDIFY AND STABILIZE
 CC THEIR MORPHOLOGY.

CC -1- DOMAIN: THE BASIC REGION CONTAINING THE REPEATS MAY BE RESPONSIBLE
 CC FOR THE BINDING OF MAP1A TO MICROTUBULES.
 CC -1- PTM: VARIOUS SERINE RESIDUES MAY BE PHOSPHORYLATED BY CAMP KINASE.
 CC -1- PTM: LC2 IS COEXPRESSED WITH MAP1A. IT IS A POLYPEPTIDE GENERATED
 CC FROM MAP1A BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
 CC BOTH MAP1A AND MAP1B.

CC -1- SIMILARITY: TO MAP1B.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M83196; AAB48069.1; -
 DR PIR: A43359; A43359.

KW Microtubules; Repeat; Phosphorylation.
 FT CHAIN 72465 2774 MAP1 LIGHT CHAIN LC2.
 FT DOMAIN 309 496 LYS-RICH (BASIC).
 FT REPEAT 336 541 11 X 3 AA REPEATS OF K-K-[DE].
 FT REPEAT 415 417 1.
 FT REPEAT 415 417 2.
 FT REPEAT 420 422 3.
 FT REPEAT 424 426 4.
 FT REPEAT 427 429 5.
 FT REPEAT 431 433 6.
 FT REPEAT 436 438 7.
 FT REPEAT 440 442 8.
 FT REPEAT 444 446 9.
 FT REPEAT 449 451 10.
 FT REPEAT 539 541 11.

SQ SEQUENCE 2774 AA; 299526 MW; 3DE74427BA9D7D7 CRC64;
 Query Match 6.8%; Score 92; DB 1; Length 2774;
 Best Local Similarity 22.4%; Pred. No. 12e+02;
 Matches 63; Conservative 46; Mismatches 100; Indels 72; Gaps 15;

QY 11 KKEVIRKSGLSAGKSVYFFSPGKFKPSKPOLAVYLGNTYDLSFDRTKMM----- 65
 DB 431 KKEGRKEERKKDAKDE-----KRKDTKPEVKKL--SKPDLKFPTEPVAKTLYKAKA 480
 QY 66 PSKLOKRLRLNDPLNQNKGKPDINT---TLPIROTASTIFKQPVTKYTNHPSKVKSDP 120
 DB 481 PGRYKVDKGRRA-----RGEKELSSRPRTPPAQGAA---PPAAVSGHRELALSSPE 529
 QY 121 ----DQRMNEQPRQLFWKRIQGLS---ASDYTEQI---IKTMELPKG--LQGVCP 165
 DB 530 DLQDFEELKREERGLAQRDGTGLGKPLPADATEQGHPAALIVQ--PSPVYLEGSHV 588
 QY 166 GSNDETLTLLSAVASALHTSSAPITGQSAVAKENPAVWLTSPQLCAATVDEDIRKOE 225
 DB 589 EREKEVPPSPGDKSTNGPDSG--AEVEKEKETW-----ERRKORE 629
 QY 226 ----RVQYRKRLLEALMADILSRADTTEMDIDMSGDE 261
 DB 630 AELGPENTAARESEAEVKEVDYER-AELEMEETHHPDDE 669

RESULT 14

AROQ_PSES2 STANDARD; PRT: 449 AA.

AC P56952;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-

DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPs).
 GN AROA.
 OS Pseudomonas sp. (strain PG2982), and
 OS Acetomobacter sp. (strain LBMA).
 OC Bacteria; Proteobacteria.
 NCBI_TaxID=308, 129026;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 2-16.
 RA Barry G.F., Kishore G.M., Padgett S.R., Stallings W.C.;
 RT "Glycosyltransferase 5-enolpyruvylshikimate-3-phosphate synthases";
 CC patent number US5633433, 27-MAY-1997.
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate -
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
 CC sixth step.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- MISCELLANEOUS: RESISTANT TO THE ANTIBIOTIC GLYPHOSATE.
 CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
 DR InterPro: IPR001986; EPSP_synthase.
 DR Pfam: PF00275; EPSP_synthase; 1.
 DR ProDom: PD001867; EPSP_synthase; 1.
 DR PROSITE: PS00104; EPSP_synthase; 1.
 DR PROSITE: PS00885; EPSP_synthase; 2; 1.
 KW Aromatic amino acid biosynthesis; Transferase; Herbicide resistance.
 SQ SEQUENCE 449 AA, 47297 MW, 447213ECCAFEC1 CRC64;

Query Match 6.7%; Score 90.5; DB 1; Length 449;
 Best Local Similarity 22.0%; Pred. No. 15;
 Matches 57; Conservative 38; Mismatches 123; Indels 41; Gaps 9;

QY 16 IRKSGISAGKSDVYIFSPGKFRSKPOLARYLGN-----TVDL-SFDFRTGKMPS 67
 DB 70 IRREG-----DWTITNGVNGCLLOPEALDPGNAGTCARLTMGLVGYDMKTSFIGDA 123
 QY 68 KLOKNKQRLNDPLNOKKRPDL-TTLPPIROTASIFKQPYTKYTNHPSNKKVSDPQRM 125
 DB 124 SLKRRMGRVRLNREMGVVEAADDGRMPLTLIGPKTANPLTYRVPMASADVKR----- 178
 QY 126 NEOPROLFWKRLQGLSASIVT---EQIITKTMELPRLGCVGPGSNDFTLSAVASALHT 182
 DB 179 -----AVLAGLNTPGVTIVIEPVMTRDTEKMLQGFGLDLYETDKDGVRIHIRT 229
 QY 183 SSAPITGOVSAVEKNPAWLNTSOPLCARFYTDIEDIRKQEEVVOQVRKKLEALMADI 242
 DB 230 GQGLVGO-TIIVPGDS---STAPFLVALLVGSDVTIRNVLNMPTR---TGLILTL 281
 QY 243 LSRADTEMDIEMSGDE 261
 DB 282 QEMGADIEVLNARLAGED 300

RESULT 15
 Y364_RICPR STANDARD; PRT; 386 AA.
 ID Y364_RICPR
 AC Q9ZDC5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein RP364.
 GN RP364.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 CC Rickettsiaceae; Rickettsiae; Rickettsia.
 NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sichert-Berger T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of

RT mitochondria";
 RL Nature 396:133-140(1998).
 CC -1- SIMILARITY: SOME. TO R.PROWAZEKII RP363.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC EMBL: AJ235271; CAAL4823.1; -;
 DR Hypothetical protein; Complete proteome.
 KW SEQUENCE 386 AA, 44346 MW, AAEB08B8D43A0E7 CRC64;

Query Match 6.7%; Score 90; DB 1; Length 386;
 Best Local Similarity 21.3%; Pred. No. 13;
 Matches 50; Conservative 45; Mismatches 74; Indels 66; Gaps 11;

QY 8 PGWKEEVIRKSGISAGKSDVYIFSPGKFRSKPOLARYLGNTVLSSDFRTGKMPS 67
 DB 170 PLFRESDIYKKLGLKSAEIEREIODPNGK-----YVOQLIDA-----KIGSNIFR 214
 QY 68 KLOKNKQRLNDPLNOKKRPDL-TTLPPIROTASIFKQPYTKYTNHPSNKKVSDPQRM 127
 DB 215 HMOKN-----NVNNGK-----EIERTA-IIEKATTKF--EODKFSFEGKKRDE 255
 QY 128 QPROLFWKRLQGLSASIVT---EQIITKTMELPRLGCVGPGSNDFTLSAVASALHTSAP 187
 DB 256 ITKYL--SKSLG--ASDYILTFKKNELVYIYGLDKG---QTLMSKVNANVIGIYSYST 308
 QY 188 TGOVSAVEKNPAWLNTSOPLCARFYTDIEDIRKQ-----EEVVOQVRKKL 234
 DB 309 SKENLKSVAK-----IINDKISSHTPLKIEVDKILKQISKEL 346

Search completed: March 12, 2003, 08:44:15
 Job time : 17.511 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2003, 03:22:03 ; Search time 19 0758 Seconds

(without alignments)
1320.377 Million cell updates/sec

Title: US-09-554-414b-2_COPY_150_411

Perfect score: 1344
Sequence: 1 MDCPALPPGKMKKEVIRKSG.....LSRADTREMIDEMSGDEA 262

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	177	13.2	476 2	S57963 methyl Cpg-binding
2	167	12.4	462 2	A41907 methyl-Cpg-binding
3	118	8.8	186 2	T48092 hypothetical prote
4	117.5	8.7	384 2	G86287 hypothetical prote
5	115	8.6	820 2	T12972 hypothetical prote
6	109.5	8.1	182 2	T45595 hypothetical prote
7	107.5	8.0	955 2	S52959 male-specific leth
8	105.5	7.8	527 2	B64633 hypothetical prote
9	104.5	7.8	512 2	T01769 hypothetical prote
10	104.5	7.8	1430 2	T34516 hypothetical prote
11	103	7.7	566 2	T04569 hypothetical prote
12	101	7.5	792 2	E84525 hypothetical prote
13	100	7.4	375 2	E81442 probable MCP-domai
14	99	7.4	789 2	E84514 hypothetical prote
15	99	7.4	1009 2	S49618 helicase-like tran
16	98.5	7.3	649 2	T37740 coiled-coil protei
17	97	7.2	448 2	I50730 yes-associated pro
18	97	7.2	684 2	C84434 DNA mismatch repa
19	95	7.1	794 2	T37989 Rhod-binding prote
20	93.5	7.0	1024 2	T30868 SCD25 protein (ver
21	93	6.9	1048 2	S64758 regulatory protein
22	92.5	6.9	5762 2	A41819 probable integrase
23	92	6.8	660 2	I60167 bifunctional purin
24	92	6.8	345 2	S77631 proliferation pote
25	92	6.8	493 2	G89876 microtubule-associ
26	92	6.8	1560 2	T42727 hypothetical prote
27	92	6.8	1746 2	D83181 microtubule-associ
28	92	6.8	2774 2	A43359 hypothetical prote
29	91.5	6.8	611 2	T14738

ALIGNMENTS

30	91	6.8	345 2	S77632	probable integrase
31	91	6.8	357 2	T48747	hypothetical prote
32	91	6.8	643 2	G87626	DNA primase (Impr
33	91	6.8	1250 91	S14177	SCD25 protein (ver
34	91	6.8	1610 2	D89451	protein t0469.1 [1
35	90.5	6.7	469 2	T46929	hypothetical prote
36	90.5	6.7	459 2	T46930	hypothetical prote
37	90.5	6.7	496 2	T15691	hypothetical prote
38	90.5	6.7	622 2	C84506	probable Achilla re
39	90.5	6.7	658 2	E71465	hypothetical prote
40	90.5	6.7	934 2	T47546	hypothetical prote
41	90.5	6.7	990 2	I51618	protein kinase-lik
42	90.5	6.7	1166 2	T13958	nucleolar phosphop
43	90.5	6.7	1249 2	T14270	syncAP-b1 protein
44	90.5	6.7	1293 2	T14259	Ras-GTPase activat
45	90	6.7	345 2	A89991	ras GTPase-activat
					Integrase (Importe

RESULT 1

S57963 methyl Cpg binding protein 2 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 05-Nov-1999

C/Accession: S57963

R/d'Esposito, M.; Quaderi, N.A.; Ciccodicola, A.; Brunl, P.; Esposito, T.; D'Urso, M.

submitted to the EMBL Data Library, July 1995

A:Description: Physical mapping and expression analysis of an x-linked gene encoding

A:Reference number: S57963

A:Accession: S57963

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-476 <DE>

A:Cross-references: EMBL:X89430; NID:g899295; PION:CAA61599.1; PID:g899296

Query Match

Best Local Similarity 28.4%; Pred. No. 1e-05; Length 476;

Matches 75; Conservative 37; Mismatches 86; Indels 66; Gaps 14;

QY	2	DGPALPPGKMKKEVIRKSGLSAGKSDVYFSPGKFRKPOLARY---LGNT-VDLSEF	57
DB	86	DDPTLEGEWTRKLQKRSRSGAKYDVLLINQKAFRKSVELIAFEVGTSLDPNDF	145
QY	58	DFR-TKKMPSKLOKQKRLRNDP---LNQKGR-DLNTTLPTRQTASIFKQPYTKVT	111
DB	146	DFVTGRGSPSRREOKPPKPKSPKAPGTGRGRGPKSGTTPRPAATSEGV--VKRYL	203
QY	112	NHPSNKKVSDPQRMNEQPOLFEKRLQGLSADYTE-QIIR-----TWELPKG	159
DB	204	E-----KSPGLILVKNMPQTSPGKAEGGATTTSTQWVIRPKRKAADPOAIRK-	256
QY	160	LOGVCGSGNDETLISAVASALHTSSAPITGOVSAVERNPAYMLTSQLCKAFIVTDED	219
DB	257	KGGRKRG-----SVVAAA-----AAEAKKA-----VKSSS	282
QY	220	IRKQERVOQVRK-KLEELMADI	242
DB	283	IRSVQETVLPKRRKTRTSTEV	306

RESULT 2

A41907 methyl-Cpg-binding protein 2 - rat

N:Alternate names: chromosomal protein Mecp2

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999

C/Accession: A41907; S41461

R:Lewis, J.D.; Meehan, R.R.; Henzel, W.J.; Maurer-Fogy, I.; Jepsen, P.; Klein, F.;

Cell 69, 905-914, 1992

A:Title: Purification, sequence, and cellular localization of a novel chromosomal pro

A:Reference number: A41907; MUID:92298389; PMID:1606614

A:Accession: A41907
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-492 <LEM>
A:Cross-references: GB:M94064; NID:g205361; PID:AAA41584.1; PID:g205362
R:Man, X.; Meenan, R.R.; Bird, A.
Nucleic Acids Res. 21, 4886-4892, 1993
A:Title: Dissection of the methyl-CPG binding domain from the chromosomal protein MeCP2.
A:Reference number: S41461; MUID:94232813; PMID:8177735
A:Contents: annotation; methyl CPG-binding domain
C:Keywords: chromosomal protein; DNA binding
F:78-162/Domain: methyl-CPG-binding #status experimental <MCG>

Query Match 12.4%; Score 167; DB 2; Length 492;
Best Local Similarity 28.1%; Pred. No. 5.9e-05;
Matches 74; Conservative 37; Mismatches 88; Indels 64; Gaps 13;

QY 2 DCPALPFGMKKEVIRKSGLSAGKSDVYFSPGKFRSKPOLARYL---LGNT-VDLSSF 57
DB 96 DDPTLPFGWTRKLRKQKSGSAGKSDVYVILNPGKAFKSKVELIAFEKVGDTSLDPNDF 155
QY 58 DFR-TGKMPKSLQKNKORLNDP---LNQNGKPDLTLPPIROTASIFKQPYTKYTN 112
DB 156 DFTYTGSPSRREKQPKKSPKAPGTCGRGREGSGTGPRKAASEGVO-VKRVLE 214
QY 113 HPSNKKVSDPQRMNEOPRQLFWEKRIQGLASAVTE-QIILK-----TMELPKGL 160
DB 215 -----KSPKLIYKMPFQSPGSGEGGATTSAGVWYIKRGRKRAKADPQALPK-K 267
QY 161 QGVGPGSNDFTLSAVASALHTSSAPITGQVSAAVEKNPAVWLNTSQPLCKAFIYDEDI 220
DB 268 RGRKPG-----SVAAA-----AAEAKKKA-----YKESSI 293
QY 221 RKOERYOVQR-KLEALMADI 242
DB 294 RSVQETVLPKIKRKTRETSLEV 316

RESULT 3
T48092
hypothetical protein T20010.130 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48092
R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Lem
submitted to the Protein Sequence Database, April 2000
A:Reference number: 224484
A:Accession: T48092
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <OHE>
A:Cross-references: EMBL:AL163816
A:Experimental source: cultivar Columbia; BAC clone T20010
C:Genetics:
A:Map position: 3
A:Introns: 16/3
A:Note: T20010.130

Query Match 8.8%; Score 118; DB 2; Length 186;
Best Local Similarity 28.8%; Pred. No. 0.08; 38; Indels 18; Gaps 3;
Matches 30; Conservative 18; Mismatches 88;

QY 4 PALPFGMKKEVIRKSGLSAGKSDVYFSPGKFRSKPOLARYL-----NTVDLSSFD 58
DB 91 PFTPGKFSKSLVLR---DYKMDTYITPTGKLRSENEIAAIVANPEFRNAPLGDFN 147
QY 59 FRTGKMPKSLQKNKORLNDP---LNQNGKPDLTLPPIROTASIFKQPYTKYTN 102
DB 148 FTVPRVMEDEV-----PPDKLGSPPSTTTTSSKSSV 181

RESULT 4
G86287

hypothetical protein F9L1.28 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: G86287
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conon, L.; Conway, A.B.; Conway, A.R.; Greasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maltl, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86287
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <STO>
A:Cross-references: GB:AE005172; NID:g5103831; PID:AMD39661.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 8.7%; Score 117.5; DB 2; Length 384;
Best Local Similarity 20.2%; Pred. No. 0.23;
Matches 60; Conservative 61; Mismatches 137; Indels 39; Gaps 13;

QY 1 MDCPALPFGMKKEVIRKSGLSAGKSDVYFSPGKFRSKPOLARYL-----GNTVDLSS 56
DB 10 TELPA-PASWKKLFYPRKAG-TRKTEIVFVAPTEGELISRRQDLQYLAHGNV-ISE 66
QY 57 FDFRTGKMPKSLQKNKORLNDP-----NDPL-----NONGKPDLTLPPIROT 99
DB 67 FEWTTGE-TPRRSSRIQKVKATPTPDKEPLTKRRSSLTFRKDNKAEEKNEAAREN 125
QY 100 ASIFKQPYTVYTHNPSKVK-----SDPQRMNEOPRQLFWEKRIQGLASAVTEQIILK 152
DB 126 MDVADKDKTEVAEAKKEKEGVTETAEAKENBEKTEAEKVNKEGTAGKEGOTE 185
QY 153 TMELPKGLQGVGSGNDFT-LTSAVASALHTSSAPITGQVSAAVEKNPAVWLNTSQ 207
DB 186 IAEKKEKEKEKAENKAEVVRDKESMEVDTSELEKAGSGEGCAEPSPVEGLDTE 245
QY 208 PLCKAFIYDEDIRKQ--EERVQOVRRKLEALMADILSRADTEMDIEMSGDEA 262
DB 246 MKEQEVVTEADVEKKEPEKTEKSGSVTTEANGQNVTLGEPNLDADAADKRES 302

RESULT 5
T12972
hypothetical protein T2118.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 18-Aug-2000
C:Accession: T12972
R:Choisme, N.; Robert, C.; Brothier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.;
submitted to the Protein Sequence Database, July 1999
A:Reference number: Z17586
A:Accession: T12972
A:Molecule type: DNA
A:Residues: 1-820 <CHO>
A:Cross-references: EMBL:AL096860; GSPDB:GN00061; ATSP:T2118.10
A:Experimental source: cultivar Columbia; BAC clone T2118
C:Genetics:
A:Gene: ATSP:T2118.10
A:Map position: 3
A:Introns: 115/1; 297/3; 431/3; 443/3; 484/3; 520/3; 539/3; 560/3; 641/2; 668/3; 726/
C:Superfamily: Arabidopsis thaliana hypothetical protein T2118.10

Query Match 8.6%; Score 115; DB 2; Length 820;
Best Local Similarity 23.1%; Pred. No. 0.97; 99; Indels 80; Gaps 14;
Matches 65; Conservative 37; Mismatches 99;

QY 18 KSLGAGKSDVYFSPS---GKFRSKPOLARYLGNVTDLSSFDFTGKMPKSLQK--- 71

A: molecule type: DNA
A: Residues: 1-955 <PAL>
A: Cross-references: EMBL:LI4582
R: Palmer, M.J.; Megner, V.A.; Richman, R.; Manning, J.E.; Kuroda, M.I.; Lucchesi, J.C.
submitted to the EMBL Data Library, April 1993

RESULT 9
T01769
Hypothetical protein A_IG002P16.18 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 15-Sep-2000
C:Accession: T01769

QY 72 -NKORLND-----PLNOKGK-----PDLNTLPIROTASIFKOPVTNHNPSNKKY 119
 Db 543 NDCGNDNDNDNDQFQAPGRKSKRQTVPSIHQAPE--TAEKKKHPII-----HPRAKV- 595
 QY 120 SDPORMNEQROFLFWERKLO-----GLSADVTEQIITKIMELPKGLGVGPGSNDLTLISA 175
 Db 596 -DNRLEKLAEE--WKSRRKKNPLSLAGNIVDKMFTTLETP-----GK 635
 QY 176 VASALHTSSAPITGOVSAAYE-----KNPAV-----WLTNSQPLCKAFIYT 216
 Db 636 AITATH-----VDALIELMKTRKESNPFLFKKSVFVGVSSFLNVIDESYMEFLDN 686
 QY 217 DEIRKOEERVOQVRKKLEEA--LMADILSRAADTEMDI 254
 Db 687 KEGFOFOSEIEKEVKKKRYAAVSITPPIVIRNLKKEDMDV 727

RESULT 13

E81442
 Probable MCP-domain signal transduction protein Cj0246c [Imported] - Campylobacter jejuni
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C:Accession: E81442
 R:Parthill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rejandream, M.A.; Rutherford, K.M.; Vanyile, A.; Whitehead, S.; Barrell
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
 A:Reference number: A81250; MUID:20150912; PMID:10688204
 A:Accession: E81442
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-375 <PAR>
 A:Cross-references: GB:AL139074; GB:AL111168; NID:96967505; PIDN:CAB72714.1; PID:9696772
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj0246c

Query Match 7.4%; Score 100; DB 2; Length 375;
 Best Local Similarity 20.6%; Pred. No. 4.6;
 Matches 54; Conservative 47; Mismatches 101; Indels 60; Gaps 10;

QY 20 GLSACKSDVYFFSPSGKKFRS--KPOLARYLGNITVD--LSDFDRGKMKPKLQKNO--R 75
 Db 105 GVGIAKSLVAFEFVEVENFNFR--RYLFTIQKISDRKYRYLNRQKLEMTNEINK 161
 QY 76 LRNPDLNOKKRPDLNTTL-----PIRQTASIFKOPVTNHNPSNKKVSDPORNN 126
 Db 162 IILNQLKQNTSEIDKNTSIFKEIPELEENIRRSITTSSEFTNFTVLEYNLQSEKKN 221
 QY 127 EQPRLFEKRLQGLSADVTEQIITKIMELPKGLGVGPGSNDLTLISAVASALHTSSAP 186
 Db 222 -----LEKRYQ--SSLKNDIDNLSLTLISGIA-----EQTILSLMGIEAARAG 265
 QY 187 ITGOVSAAVEKNPAWLNLTSQLCAFIYVDEDIRKOEERVOQVRKKLEALMADI----- 242
 Db 266 KLG-----RGFAVYADDEVRLKLENTQMGELGEMGAIKVIQTITQ 304
 QY 243 ---LSRAADTEMDIEMDSGE 261
 Db 305 SIAKSSNSTQEMNFTROKTN 326

RESULT 14

E84514
 Hypothetical protein Atg14130 [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: E84514
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.;
 Narens, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: E84514
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1789 <STO>
 A:Cross-references: GB:AE002093; NID:94587682; PIDN:AAD25852.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: Atg14130
 A:Map position: 2
 C:Superfamily: Arabidopsis thaliana hypothetical protein T2118.10

Query Match 7.4%; Score 99; DB 2; Length 789;
 Best Local Similarity 21.9%; Pred. No. 15;
 Matches 59; Conservative 38; Mismatches 96; Indels 76; Gaps 12;

QY 12 KEVIRKSGLSAGSDVYFFSPSGKKFRS--KPOLARYLGNITVDLSDFDRGKMKPKLQK 71
 Db 506 RKQVPRKROKQVDSADVDV--PTRKEAKSKR--KITGNDNDNDNDQFQAPGRK 561
 QY 72 NKORLNDPLNOKKRPDLNTLPIROTASIFKOPVTNHNPSNKKVSDPORNEQPRQ 131
 Db 562 SKRGT-----VPSIHQAPE--TAEKKKHPII-----HPRAKV--DNRLEKLAEE- 602
 QY 132 LMEKRLQ-----GLSADVTEQIITKIMELPKGLGVGPGSNDLTLISAVASALHTSSAPI 187
 Db 603 --WKSRRKKNPLSLAGNIVDKMFTTLETP-----GKAIT 635
 QY 188 TGOVSAAYE-----KNPAV-----WLTNSQPLCKAFIYVDEDIRKOEERVO 228
 Db 636 TTHVDALELMKTRKESNPFLFKKSVFVGVSSFLNVIDESYMEFLDNKGFQFOSEIEK 695
 QY 229 QVRKLEEA--LMADILSRAADTEMDI 254
 Db 696 EVKRRYAAVSITPPIVIRNLKKEDMDV 724

RESULT 15

S49618
 Helicase-like transcription factor - human

C:Species: Homo sapiens (man)
 C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 02-Aug-2002
 C:Accession: S49618; A56095
 R:ding, H.; Descheememaker, K.; Marynen, P.; Nelles, L.; Carvalho, T.; Carmo-Fonseca,
 submitted to the EMBL data library, November 1994
 A:Description: Characterization of a helicase-like transcription factor involved in p
 A:Reference number: S49618
 A:Accession: S49618
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1009 <DIN>
 A:Cross-references: EMBL:246606; NID:9575250; PIDN:CAA86571.1; PID:9575251
 R:Sheridan, P.L.; Schorpp, M.; Voz, M.L.; Jones, K.A.
 J. Biol. Chem. 270, 4575-4587, 1995
 A:Title: Cloning of an SNF2/SWI2-related protein that binds specifically to the SPB
 A:Reference number: A56095; MUID:95181452; PMID:7876228
 A:Accession: A56095
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 134, 'P', 36-336, 'D', 338-381, 'T', 383-912, 'K', 914-1009 <SHE>
 A:Cross-references: GB:L34673; NID:9531195; PIDN:AAA67436.1; PID:9531196
 C:Genetics:
 A:Gene: HLT-1; GDB:SNF2L3; HIP116
 A:Cross-references: GDB:392575
 A:Map position: 3925.1-3926.1
 C:Superfamily: human PMR-1 protein; RING finger homology
 C:Keywords: zinc
 P:756-806/Domain: RING finger homology <RNG>

Query Match 7.4%; Score 99; DB 2; Length 1009;
 Best Local Similarity 22.3%; Pred. No. 21;
 Matches 57; Conservative 50; Mismatches 79; Indels 70; Gaps 15;

Mon Mar 17 08:44:16 2003

us-09-554-414b-2_copy_150_411.rpr

Page 6

[illegible]

Search completed: March 12, 2003, 09:13:57
Job time : 26.0758 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2003, 05:39:48 ; Search time 12.0684 Seconds
(without alignments)
915.501 Million cell updates/sec

Title: US-09-554-414b-2_COPY_150_411
Perfect score: 1344
Sequence: 1 MDCPALPGMKKEVIRKSG.....LSRADTEPMIDEMSGDEA 262

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues
Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	411	10	US-09-749-728B-1
2	104	7.7	433	9	US-09-906-514-2
3	100	7.4	2568	10	US-09-866-108-3
4	91.5	6.8	608	10	US-09-799-777-7
5	90.5	6.7	449	9	US-09-464-099A-5
6	90.5	6.7	449	9	US-09-464-099A-7
7	90.5	6.7	449	10	US-09-861-696-5
8	90.5	6.7	449	10	US-09-861-696-7
9	90.5	6.7	497	10	US-09-764-864-1314
10	90.5	6.7	534	10	US-09-764-864-861
11	90.5	6.7	1111	10	US-09-815-242-12955
12	90	6.7	2665	10	US-09-864-761-34248
13	89	6.6	507	10	US-09-876-187-2
14	89	6.6	507	10	US-09-749-728B-13
15	88.5	6.6	428	9	US-09-906-514-4
16	88.5	6.6	660	9	US-09-872-462-4
17	88.5	6.6	939	10	US-09-226-248B-29
18	88.5	6.6	939	10	US-09-801-368-380
19	88	6.5	465	10	US-09-749-728B-17

20	87	6.5	2285	10	US-09-932-183A-2	Sequence 2, Appl1
21	86.5	6.4	459	10	US-09-968-915-1	Sequence 1, Appl1
22	86	6.4	478	9	US-10-106-534-6	Sequence 6, Appl1
23	86	6.4	413	9	US-09-925-299-903	Sequence 903, Appl
24	86	6.4	478	10	US-09-925-299-903	Sequence 903, Appl
25	86	6.4	26926	9	US-09-759-508B-2	Sequence 2, Appl1
26	85.5	6.4	1404	10	US-09-811-945A-1	Sequence 1, Appl1
27	84.5	6.3	229	10	US-09-968-915-7	Sequence 7, Appl1
28	84.5	6.3	1043	10	US-09-946-805-4	Sequence 4, Appl1
29	84	6.2	473	10	US-09-876-187-6	Sequence 6, Appl1
30	84	6.2	473	10	US-09-876-187-8	Sequence 8, Appl1
31	84	6.2	1093	10	US-09-801-368-392	Sequence 392, Appl
32	83	6.2	96	10	US-09-864-761-37960	Sequence 37960, A
33	83	6.2	96	10	US-09-864-761-36987	Sequence 36987, A
34	83	6.2	722	10	US-09-815-242-12888	Sequence 12888, A
35	83	6.2	792	9	US-10-025-380-1127	Sequence 1127, Ap
36	83	6.2	945	10	US-09-745-763-191	Sequence 191, Ap
37	83	6.2	991	10	US-09-815-242-5803	Sequence 5803, Ap
38	83	6.2	1711	10	US-09-771-161A-219	Sequence 219, App
39	83	6.2	1711	10	US-09-771-161A-220	Sequence 220, App
40	82.5	6.1	578	10	US-09-159-469-50	Sequence 50, Appl
41	82.5	6.1	578	10	US-09-798-042-50	Sequence 4, Appl1
42	82.5	6.1	654	10	US-09-940-921B-4	Sequence 11396, A
43	82	6.1	858	10	US-09-815-242-11396	Sequence 8, Appl1
44	82	6.1	1122	9	US-10-072-094-81	Sequence 91, Appl
45	82	6.1	1122	9	US-10-072-094-91	

ALIGNMENTS

RESULT 1
US-09-749-728B-1
: Sequence 1, Application US/09749728B
: Patent No. US20020142457A1
: GENERAL INFORMATION:
: APPLICANT: Umezawa, Akihito
: APPLICANT: Hata, Jun-ichi
: APPLICANT: Fukuda, Keiichi
: APPLICANT: Ogawa, Satoshi
: APPLICANT: Sakurada, Kazuhiko
: APPLICANT: Gojo, Satoshi
: APPLICANT: Yamada, Yoji
: TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDI
: FILE REFERENCE: 00766.000043
: CURRENT APPLICATION NUMBER: US/09/749,728B
: CURRENT FILING DATE: 2001-09-17
: PRIOR APPLICATION NUMBER: H11-372826
: PRIOR FILING DATE: 1999-12-28
: PRIOR APPLICATION NUMBER: PCT-JP00-01148
: PRIOR FILING DATE: 2000-02-28
: PRIOR APPLICATION NUMBER: PCT-JP00-07741
: PRIOR FILING DATE: 2000-11-02
: NUMBER OF SEQ ID NOS: 80
: SOFTWARE: PatentIn Ver.2.0
: SEQ ID NO 1
: LENGTH: 411
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-749-728B-1
Query Match 100.0%: Score 1344: DB 10: Length 411:
Best Local Similarity 100.0%: Pred. No. 2.7e-113:
Matches 262: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 MDCPALPGMKKEVIRKSGISACKSDVYFSPSGKFRSKPOLARYLGTVDLSFDPR 60
DB 150 MDCPALPGMKKEVIRKSGISACKSDVYFSPSGKFRSKPOLARYLGTVDLSFDPR 209
QY 61 TGKMPKSLKNNKORLNDPLNKKGKPDNTLPIQOTASITKOPYTKYTNHPSKVKVS 120
DB 210 TGKMPKSLKNNKORLNDPLNKKGKPDNTLPIQOTASITKOPYTKYTNHPSKVKVS 269

QY 121 DPGANQOPQLEWKRLOGLSADVTQIITKMTLPKGLQGVGSGNDITLLSAYASAL 180
Db 270 DPGANQOPQLEWKRLOGLSADVTQIITKMTLPKGLQGVGSGNDITLLSAYASAL 329
QY 181 HTSSAPITGGVSAVEKNPVMNTSOPICAFIVTDEDIRKOEERVOQVRKKLEBALMA 240
Db 330 HTSSAPITGGVSAVEKNPVMNTSOPICAFIVTDEDIRKOEERVOQVRKKLEBALMA 389
QY 241 DILSRADTEEMDIEMDSGDEA 262
Db 390 DILSRADTEEMDIEMDSGDEA 411

RESULT 2
US-09-514-2
; Sequence 2, Application US/09906514
; Patent No. US20020170085A1
; GENERAL INFORMATION:
; APPLICANT: Kaeppler, Shawn
; APPLICANT: Springer, Nathan
; APPLICANT: Phillips, Ronald
; TITLE OF INVENTION: Methyl Cpg Binding Domain Nucleic Acids from Maize
; FILE REFERENCE: Methylation
; CURRENT APPLICATION NUMBER: US/09/906,514
; CURRENT FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Zea mays
US-09-514-2

Query Match 7.7%; Score 104; DB 9; Length 433;
Best Local Similarity 19.9%; Pred. No. 0.15;
Matches 53; Conservative 48; Mismatches 99; Indels 66; Gaps 11;

QY 4 PALPQKKEEVIRKSLGSKSDYVYFSPSGKKFRKPOLAKYL-----GNTVDLSFPD 58
Db 2 PA-PDGVTK-----KFTPLRGGRSEIVSPTEGEEKRRLSOLYLAHNGPAPV--SEFD 54
QY 59 FRGKMPKSKLQKKNKRLNDPLNQNKGKPDLTTLPIROTASIFKOP---VTKYVNH 114
Db 55 WGTG-----DTPRSAR-----ISEKAVFSDPBEKIPKRSRNS 89
QY 115 SNK-----VSDPGRMEQOPQLEWKRLOGLSADVTQIITKMTLPKGLQGVGSGND 169
Db 90 SGRGKQGGKKEPPE--TEAKDAETGKEAEAPSEDAKEDVDMKPAEEVKGASAEED 147
QY 170 ETLISAVASALHTSSAPITGGVSAVEKNPVM-----LNTSOPICAFIVTDEDIRKOE 224
Db 148 ADMADAPAPAPMEEDKQTELEALAIAPVPSEKKDVPAPPEAASNPTEDSAPAPA 207
QY 225 ERYQOVRRKLEALMADILSRADTE 250
Db 208 E-----PADVAPAPAEATK 220

RESULT 3
US-09-866-108-3
; Sequence 3, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: A60MICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108

;; CURRENT FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 60/266,860
;; PRIOR FILING DATE: 2001-02-05
;; NUMBER OF SEQ ID NOS: 15752
;; SOFTWARE: A60MICA Sequence Listing Engine
;; SEQ ID NO 3
;; LENGTH: 2568
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-866-108-3

Query Match 7.4%; Score 100; DB 10; Length 2568;
Best Local Similarity 23.7%; Pred. No. 4.1;
Matches 42; Conservative 24; Mismatches 49; Indels 62; Gaps 7;

QY 139 OGLSADVTQIITKMTLPKGLQGVGSGNDITLLSAYASALHTSSAPITGG----- 190
Db 1290 EGIDERNKAELEITLDELK--KAVAGHSQVFLAGVISRLEKREKLVSQIVLFGAA 1347
QY 191 -----VSAAEKNPAP-----W--LNTSOPICAFIVTDEDIR 221
Db 1348 CKGFLSRQEFKKIKIRLAQCIOKNVAVFLAVKDPMPWQQLGSIQPLISATIGT-EOLR 1406
QY 222 KQERQOQVRKLE-----EALMADILSRADTEEMDIEMDSGDEA 262
Db 1407 AKRELLTLRKLKSEKRLNELRONTDLLESKIADLTSLAD-----ERKGGVA 1457

RESULT 4
US-09-799-777-7
; Sequence 7, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hallman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baugh, Marian
; APPLICANT: Sather, Susan
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSER: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE

CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/799,777
FILING DATE: 06-Mar-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/002,485
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0459 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGFET03
CLONE: 1255027
SEQUENCE DESCRIPTION: SEQ ID NO: 7 :
US-09-799-777-7
Query Match 6.8%; Score 91.5; DB 10; Length 608;
Best Local Similarity 24.3%; Pred. No. 3.2;
Matches 57; Conservative 30; Mismatches 93; Indels 55; Gaps 10;
QY 54 LSFDFRTGKMPKSL-----QKNKRLNDPLNKKGRPDLNTITPIR 97
DB 83 LQDFDENVAKVAFVCGSAIQVKEKNMGGKKNRKRKSKS-KOHGKNDADKVERP 141
QY 98 QTASIFKQPVYKVTNHPNS--KVKSPQRMNEQPROLFEKRLQGLSASDVEQIITKM 154
DB 142 EAGPLQPP-POIONGPMNGCEKSSSTDSANEKPAIIPREKTSILE----- 188
QY 155 ELPGLOGVGGNDELILSAVASALHTSSAPITG-----QVSAVEKNPA----- 200
DB 189 EPSKALGVTEGNN--LLOOKLS-LDGNPKPIHGTTERSDGLQWSAEQPCNSKPKAKT 244
QY 201 --VWLNTSOPLCAFIYTDIEDIRKQEEVQOVRRKLEALMADILSRADTEEMD 253
DB 245 SPVKSNT--PAHLKIKPDELAKKRGNIKSVKDLQRCIVSLTRYVMIKEEVD 297
RESULT 5
US-09-464-099A-5
Sequence 5, Application US/09464099A
Patent No. US20020168680A1
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOPLPYRVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
FILE REFERENCE: 11899 0175 CUS01 MORT:175-2
CURRENT APPLICATION NUMBER: US/09/464,099A
CURRENT FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: US 09/137,440
PRIOR FILING DATE: 1998-08-20
PRIOR APPLICATION NUMBER: US 08/833,485
PRIOR FILING DATE: 1997-04-07

PRIOR APPLICATION NUMBER: US 08/306,063
PRIOR FILING DATE: 1994-09-13
PRIOR APPLICATION NUMBER: US 07/749,611
PRIOR FILING DATE: 1991-08-28
PRIOR APPLICATION NUMBER: US 07/576,537
PRIOR FILING DATE: 1990-08-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 449
TYPE: PRT
ORGANISM: Agrobacterium sp.
US-09-464-099A-5
Query Match 6.7%; Score 90.5; DB 9; Length 449;
Best Local Similarity 22.0%; Pred. No. 2.5;
Matches 57; Conservative 38; Mismatches 123; Indels 41; Gaps 9;
QY 16 IRKSGLSAGSDVYFSPSGKFRSKPQLARYLGN-----TVDL-SFDFRTGKMPK 67
DB 70 IRREG-----DWIINGVNGCCLLPQEAALDFGNAGTGARLTMGLVGYDMKTSFIGDA 123
QY 68 KLOKNNQRLNDPLNKKGRPDL--NTLPIRQTASIFQPVYKVTNHPNSKNYKSDPQRM 125
DB 124 SLKRRPMGRVLANLREMGVQVEADDDRMPLTILIGKTNPIYRVPMASAOVKS----- 178
QY 126 NEQPROLFEKRLQGLSASDVT--EQIITKMLPGLOGVGGNDELILSAVASALHT 182
DB 179 -----AVLLAGLNTPGVTIVIEPVMTROHTEKMLGFGADLVETDKGVRIIRIT 229
QY 183 SSAPITGVSAVEKNPAVWLNTSOPLCAFIYTDIEDIRKQEEVQOVRRKLEALMADI 242
DB 230 GQGLKVGQ-TIDVPGDS--STAFPLVALLEGGSDVTIRVLMNPTF---TGILTL 281
QY 243 LSRADTEEMDIEEMDSDE 261
DB 282 QEMGADIEVLNRLAGGED 300
RESULT 6
US-09-464-099A-7
Sequence 7, Application US/09464099A
Patent No. US20020168680A1
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOPLPYRVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
FILE REFERENCE: 11899 0175 CUS01 MORT:175-2
CURRENT APPLICATION NUMBER: US/09/464,099A
CURRENT FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: US 09/137,440
PRIOR FILING DATE: 1998-08-20
PRIOR APPLICATION NUMBER: US 08/833,485
PRIOR FILING DATE: 1997-04-07
PRIOR APPLICATION NUMBER: US 08/306,063
PRIOR FILING DATE: 1994-09-13
PRIOR APPLICATION NUMBER: US 07/749,611
PRIOR FILING DATE: 1991-08-28
PRIOR APPLICATION NUMBER: US 07/576,537
PRIOR FILING DATE: 1990-08-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 449
TYPE: PRT
ORGANISM: Pseudomonas sp.
US-09-464-099A-7
Query Match 6.7%; Score 90.5; DB 9; Length 449;
Best Local Similarity 22.0%; Pred. No. 2.5;
Matches 57; Conservative 38; Mismatches 123; Indels 41; Gaps 9;

OY 16 IRKSGLSAGSDVYFSPGKFRSKPOLARYLGN-----IVDL-SSPDRFGKMPMS 67
||| | : : : : :
Db 70 IRKEG-----DWTIINGVNGCLLPEDALDFGNAGTGARLTMGLVGTDMKTSFIDGA 123
||| | : : : : :
OY 68 KLOKNNKORLNDPLNOKNGKPD--NTTLPTRQASIFKOPVTKVTHNPSNKVSDPQRM 125
||| | : : : : :
Db 124 SLKSRPMGRVNLPLREMGVQVEAADGRMPLTLIGPTANPIYRVPMASAOVKS----- 178
||| | : : : : :
OY 126 NEOPROLFWERKLOGLSASDVT---EQIITKTMELPKGLGVGSGNDETLISAVASALHT 182
||| | : : : : :
Db 179 -----AVLLAGLNTPGVTVIEPVMTDRHTEKMLQGFADLTIVETDKDGRHIRIT 229
||| | : : : : :
OY 183 SSAPITGVSAAVEKNPAVNLNTSOPLCARFIYTDIEDIRKOEERVOQVRKLEALMADI 242
||| | : : : : :
Db 230 GQGLVQ--TIDVPGDPS---STAFPLVALLVGSDVTIRNVLMNPTFR---TGLILTL 281
||| | : : : : :
OY 243 LSRADTEEMDIEMDSGDE 261
||| | : : : : :
Db 282 QEMGADIEVLNARLAGGED 300
||| | : : : : :

RESULT 7

US-09-861-696-5
; Sequence 5, Application US/09861696
; Patent No. US2002007053A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
; FILE REFERENCE: 11899.0175.CNUS04 MOBT:175-4
; CURRENT APPLICATION NUMBER: US/09/861,696
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/137,440
; PRIOR FILING DATE: 1998-08-20
; PRIOR APPLICATION NUMBER: US 08/833,485
; PRIOR FILING DATE: 1997-04-07
; PRIOR APPLICATION NUMBER: US 08/306,063
; PRIOR FILING DATE: 1994-09-13
; PRIOR APPLICATION NUMBER: US 07/749,611
; PRIOR FILING DATE: 1991-08-28
; PRIOR APPLICATION NUMBER: US 07/576,537
; PRIOR FILING DATE: 1990-08-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Agrobacterium sp.
US-09-861-696-5

Query Match 6.7%; Score 90.5; DB 10; Length 449;
Best Local Similarity 22.0%; Pred. No. 2.5;
Matches 57; Conservative 38; Mismatches 123; Indels 41; Gaps 9;

OY 16 IRKSGLSAGSDVYFSPGKFRSKPOLARYLGN-----TVDL-SSPDRFGKMPMS 67
||| | : : : : :
Db 70 IRKEG-----DWTIINGVNGCLLPEDALDFGNAGTGARLTMGLVGTDMKTSFIDGA 123
||| | : : : : :
OY 68 KLOKNNKORLNDPLNOKNGKPD--NTTLPTRQASIFKOPVTKVTHNPSNKVSDPQRM 125
||| | : : : : :
Db 124 SLKSRPMGRVNLPLREMGVQVEAADGRMPLTLIGPTANPIYRVPMASAOVKS----- 178
||| | : : : : :
OY 126 NEOPROLFWERKLOGLSASDVT---EQIITKTMELPKGLGVGSGNDETLISAVASALHT 182
||| | : : : : :
Db 179 -----AVLLAGLNTPGVTVIEPVMTDRHTEKMLQGFADLTIVETDKDGRHIRIT 229
||| | : : : : :
OY 183 SSAPITGVSAAVEKNPAVNLNTSOPLCARFIYTDIEDIRKOEERVOQVRKLEALMADI 242
||| | : : : : :
Db 230 GQGLVQ--TIDVPGDPS---STAFPLVALLVGSDVTIRNVLMNPTFR---TGLILTL 281
||| | : : : : :

OY 243 LSRADTEEMDIEMDSGDE 261
||| | : : : : :
Db 282 QEMGADIEVLNARLAGGED 300
||| | : : : : :

RESULT 8

US-09-861-696-7
; Sequence 7, Application US/09861696
; Patent No. US2002007053A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
; FILE REFERENCE: 11899.0175.CNUS04 MOBT:175-4
; CURRENT APPLICATION NUMBER: US/09/861,696
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/137,440
; PRIOR FILING DATE: 1998-08-20
; PRIOR APPLICATION NUMBER: US 08/833,485
; PRIOR FILING DATE: 1997-04-07
; PRIOR APPLICATION NUMBER: US 08/306,063
; PRIOR FILING DATE: 1994-09-13
; PRIOR APPLICATION NUMBER: US 07/749,611
; PRIOR FILING DATE: 1991-08-28
; PRIOR APPLICATION NUMBER: US 07/576,537
; PRIOR FILING DATE: 1990-08-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Pseudomonas sp.
US-09-861-696-7

Query Match 6.7%; Score 90.5; DB 10; Length 449;
Best Local Similarity 22.0%; Pred. No. 2.5;
Matches 57; Conservative 38; Mismatches 123; Indels 41; Gaps 9;

OY 16 IRKSGLSAGSDVYFSPGKFRSKPOLARYLGN-----TVDL-SSPDRFGKMPMS 67
||| | : : : : :
Db 70 IRKEG-----DWTIINGVNGCLLPEDALDFGNAGTGARLTMGLVGTDMKTSFIDGA 123
||| | : : : : :
OY 68 KLOKNNKORLNDPLNOKNGKPD--NTTLPTRQASIFKOPVTKVTHNPSNKVSDPQRM 125
||| | : : : : :
Db 124 SLKSRPMGRVNLPLREMGVQVEAADGRMPLTLIGPTANPIYRVPMASAOVKS----- 178
||| | : : : : :
OY 126 NEOPROLFWERKLOGLSASDVT---EQIITKTMELPKGLGVGSGNDETLISAVASALHT 182
||| | : : : : :
Db 179 -----AVLLAGLNTPGVTVIEPVMTDRHTEKMLQGFADLTIVETDKDGRHIRIT 229
||| | : : : : :
OY 183 SSAPITGVSAAVEKNPAVNLNTSOPLCARFIYTDIEDIRKOEERVOQVRKLEALMADI 242
||| | : : : : :
Db 230 GQGLVQ--TIDVPGDPS---STAFPLVALLVGSDVTIRNVLMNPTFR---TGLILTL 281
||| | : : : : :
OY 243 LSRADTEEMDIEMDSGDE 261
||| | : : : : :
Db 282 QEMGADIEVLNARLAGGED 300
||| | : : : : :
RESULT 9
US-09-764-864-1314
; Sequence 1314, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792

SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1314
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1314

Query Match 6.7% Score 90.5; DB 10; Length 497;
Best Local Similarity 19.5%; Pred. No. 2.9;
Matches 59; Conservative 46; Mismatches 97; Indels 101; Gaps 13;

QY 7 PPGMKKEEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLNTVDLSFDRRTGKMP 66
DB 109 PPGMKKEELENSG-----LALYDK-----DGTDETEV 138
QY 67 SKLOKKN-----QRLNDP---LNQNGKRPD---LNTTLPIR---QTASIFKQPTKVTN 112
DB 139 GEIOONKSVTYDLKLVNYPGFNISTPRGIDEMRIGSPMGACQOKVFNAYLT--SN 196
QY 113 HPSKVKSDPQRMN-----QPROLFWEKRLQGLSASDVTEDQITKTMELPKGLQ----- 161
DB 197 FOAPGVKSGNKRSSHSSPGSPKK--QKNESNAGSPADMEIDSDMEVPHGSQSESFQ 253
QY 162 -----GVPGSNDETLISAVASALHTSSAPITGVSAVAKNPVWLNTSQ 207
DB 254 FQPLPDPPTPLPRGTPTPTVTPPLPKG-TPLTPSDSPQTRASGAVDED----- 303
QY 208 PLCAFIYTDIDIRKQERVOQVRKLEEA-----LMADILSRAADTEEMDIE 256
DB 304 -----ALTELEEQORRIWALEQAEVNSDSDVPVDPPLGNSVASSPCNELDLFV 357
QY 257 DSG 259
DB 358 PEG 360

RESULT 10
US-09-764-864-861
; Sequence 861, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIORITY FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 861
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-861

Query Match 6.7% Score 90.5; DB 10; Length 534;
Best Local Similarity 19.5%; Pred. No. 3.2;
Matches 59; Conservative 46; Mismatches 97; Indels 101; Gaps 13;

QY 7 PPGMKKEEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLNTVDLSFDRRTGKMP 66
DB 146 PPGMKKEELENSG-----LALYDK-----DGTDETEV 175
QY 67 SKLOKKN-----QRLNDP---LNQNGKRPD---LNTTLPIR---QTASIFKQPTKVTN 112
DB 176 GEIOONKSVTYDLKLVNYPGFNISTPRGIDEMRIGSPMGACQOKVFNAYLT--SN 233
QY 113 HPSKVKSDPQRMN-----QPROLFWEKRLQGLSASDVTEDQITKTMELPKGLQ----- 161

DB 234 FOAPGVKSGNKRSSHSSPGSPKK--QKNESNAGSPADMEIDSDMEVPHGSQSESFQ 290
QY 162 -----GVPGSNDETLISAVASALHTSSAPITGVSAVAKNPVWLNTSQ 207
DB 291 FQPLPDPPTPLPRGTPTPTVTPPLPKG-TPLTPSDSPQTRASGAVDED----- 340
QY 208 PLCAFIYTDIDIRKQERVOQVRKLEEA-----LMADILSRAADTEEMDIE 256
DB 341 -----ALTELEEQORRIWALEQAEVNSDSDVPVDPPLGNSVASSPCNELDLFV 394
QY 257 DSG 259
DB 395 PEG 397

RESULT 11
US-09-815-242-12955
; Sequence 12955, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA, 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIORITY FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIORITY FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12955
; LENGTH: 1111
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12955

Query Match 6.7% Score 90.5; DB 10; Length 1111;
Best Local Similarity 21.2%; Pred. No. 9.1;
Matches 55; Conservative 34; Mismatches 66; Indels 105; Gaps 12;

QY 25 KSDVYFSPSG-KKFRSKPOLARYLGN-----TVDLSFDRRT-----GKMP 66
DB 137 KPQTYVSAINGIEKTEHKTKHNTNHRAKUSTPDYHKESEFKSEVSAIFGTMKP 196
QY 67 SKLOKKNQRLNDP---LNQNGKRPD---LNTTLPIR---QTASIFKQPTKVTNHPNSKVKSDPQRMN 126
DB 197 KRLNG-----RIPVSK-----PSEKVESDQKXKD 221
QY 127 E-----QPROLFWEKRLQGLSASDVTEDQITKTMELPKGLQGVGSGNDETLISAVAS 178
DB 222 KIVAKTQTSQNKQLEQKQ-----NDSVVK-----QGTASKSSDENVSSY--- 261
QY 179 ALHTSSAPITGVSAVAKNPVWLNTSQPLCAFIYTDIDIRKQERVOQVRKLEEA 238

DB 262 ---TSMPTYSKVDNTIK-----IENIYASQIV--EEIRRRERKVLQKRRFKAL 307
 QY 239 M-----ADILSRAD 248
 DB 308 OOKREHKNEODAIQRAID 327

RESULT 12 US-09-864-761-34248

; Sequence 34248, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aeo mica-x-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; PRIOR FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263,6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 34248
 ; LENGTH: 2665
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL034555.2
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 14
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.5
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.3

; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.7
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 12
 ; OTHER INFORMATION: EST_HUMAN HIT: A0117052.1, EVALU0 0.00e+00
 ; OTHER INFORMATION: SWISSPROT HIT: P08640, EVALU0 3.00e-10
 US-09-864-761-34248

Query Match 6.7%; Score 90; DB 10; Length 2665;
 Best Local Similarity 21.3%; Pred. No. 34;
 Matches 65; Conservative 27; Mismatches 81; Indels 132; Gaps 14;

QY 14 EVIRKSG-LSAGKSDVYFFSPSGKKFRSKPOLARYLGNVTDLSSDFRTGKMPKSKLOKN 72
 DB 333 EYERKEGRLLKAKK-----HLKPPQADGVASAVDL-----EKLEAR 367
 QY 73 KORLRNDPLNOKKGRPDLTITPIROTASIF-----KOPVTK-V 110
 DB 368 KRFFADSNLAKAKOKPEVKKSSPEMEDARVLSKKOPDVSSREYLLNKGAEKKRPVKEI 427
 QY 111 TNHPSMKVKSDFQRMNEQPRQLFWFKERLQGLSASDVTEQIITWELPKGLQ-----GVGP 165
 DB 428 LKRESKIKID--RLN-----TVASPRCCQELASISVGS 459
 QY 166 GSNDFTLLSAVASALHTSSAPITTCOVSAAVEKNPAVWLNTSOPLOKAFIVTD-----ED 219
 DB 460 GSRPSSDLQARLGEAGESV-----ENQEVQSKRPI-----PSKPOLKQLOVLDDQGERED 511
 QY 220 IRKQ-----EERV-----QOVRKLEALMADILSRADT 249
 DB 512 VRKNYSLSNDEPTEPKSGQEKSHSVNTEKIIDIDHOSYRKQMEQ-----SRRKQ 564
 QY 250 EEMDI 254
 DB 565 MEMEI 569

RESULT 13
 US-09-876-187-2
 ; Sequence 2, Application US/09876187
 ; Patent No. US20020090603A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Okamoto, Shu-ichi
 ; APPLICANT: Lipston, Stuart A.
 ; TITLE OF INVENTION: Methods of Differentiating and
 ; FILE REFERENCE: P-LJ 4714
 ; CURRENT APPLICATION NUMBER: US/09/876,187
 ; PRIOR FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: US 60/209,539
 ; PRIOR FILING DATE: 2000-06-05
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 507
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-876-187-2

Query Match 6.6%; Score 89; DB 10; Length 507;
 Best Local Similarity 26.7%; Pred. No. 4.1;
 Matches 50; Conservative 19; Mismatches 50; Indels 68; Gaps 11;

QY 32 SPSSGKFF---RSKPOLARYLG-NTVYDLSDFRTGKMPKSKLOKNKORLRNDP-----L 81
 DB 223 SPVNGGFVNSRASPNLIGATGANSL-----GKVPYK-----SPPRGGGWL 264
 QY 82 NONKGRPDLTITPIROTASIFKOPVTKVTHNPSNKVKSDFQRMNEQPRQLFWFKERLQGL 141
 DB 265 GMSKSKPLRVYIP-----PSSKGMMPLESEER-----LELNTORI 301
 QY 142 SASDVTEQI-----IKTMELPKGLQV-----GPSNDEFTLLSAVASALHTSSAP----- 186
 DB 302 SSSQATQGLAPVPVSVTPPSLP--QGLVYSAMPTAVNTDYSILTSADLSALQGFNSPGML 359

Page 7

```
Search completed: March 12, 2003, 09:14:42
Job time : 16.0684 secs
```

1
2
3
4

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: March 12, 2003, 00:59:01 ; Search time 38.1516 Seconds

(Without alignments)

915.078 Million cell updates/sec

Title: US-09-554-414B-2_COPY_150_411

Sequence: 1 MDCPALPPGKKKEVIRKSG.....LSRAADTEEMDTEMDSGDR 262

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: /SID52/gcgdata/geneseq/emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	263	19 AAW74980	Human secreted pro
2	1344	100.0	281	20 AAY73829	Human prostate tum
3	1344	100.0	281	20 AAY48439	Human prostate can
4	1344	100.0	411	20 AAY14197	Human DNA demethyl
5	1344	100.0	411	22 AAB99915	Human protein sequ
6	1344	100.0	411	22 AAG64314	Human protein #1.
7	1344	100.0	411	22 AAG64844	Heart muscle cell
8	1327	98.7	414	20 AAY14199	Mouse DNA demethyl
9	1050	78.1	282	23 AAE22578	Xenopus laevis MBD
10	1047.5	77.9	303	23 AAE22577	Xenopus laevis MBD

11	1021	76.0	291	20 AAY14198	Human DNA demethyl
12	1013	75.4	200	20 AAY07107	Colon cancer assoc
13	1012.5	75.3	285	20 AAY14200	Mouse DNA demethyl
14	753	56.0	223	22 AAB92997	Human protein sequ
15	739	55.0	218	22 AAB93733	Human protein sequ
16	436.5	32.5	314	22 AAB66035	Drosophila melanog
17	428	31.8	339	23 AAE22575	Drosophila dMBD-11
18	384.5	28.6	226	22 ABB63747	Drosophila melanog
19	384.5	28.6	226	23 AAE22576	Drosophila dMBD-11
20	341	25.4	64	23 AAE22569	Human MBD2 methyl
21	280.5	20.9	63	23 AAE22570	Xenopus laevis MBD
22	214.5	16.0	574	21 AAY28900	Regulator of fibro
23	214.5	16.0	590	21 AAB21008	Human nucleic acid
24	214.5	16.0	629	21 AAY28899	Regulator of fibro
25	214.5	16.0	630	21 AAY28898	Regulator of fibro
26	209	15.6	64	23 AAE22567	Human MBD1 methyl
27	166	12.4	219	21 AAG02051	Human secreted pro
28	144.5	10.8	68	23 AAE22566	Human Mecp2 methyl
29	139.5	10.4	51	23 AAE22574	Drosophila dMBD-11
30	126	9.4	580	20 AAW74473	Human MBD1 endonuc
31	124.5	9.3	95	21 AAG50927	Arabidopsis thalia
32	124.5	9.3	149	21 AAG50926	Arabidopsis thalia
33	124.5	9.3	161	21 AAG50925	Arabidopsis thalia
34	123.5	9.2	95	21 AAG13952	Arabidopsis thalia
35	123.5	9.2	149	21 AAG13951	Arabidopsis thalia
36	123.5	9.2	161	21 AAG13950	Arabidopsis thalia
37	123.5	9.2	565	21 AAY44504	Human delta228-UV
38	120	8.9	68	23 AAE22568	Human MBD4 methyl
39	119	8.9	248	22 AAB71902	Novel human diagno
40	118	8.8	133	21 AAB24760	Plant SDF encoded
41	118	8.8	186	21 AAB24759	Plant SDF encoded
42	118	8.8	203	21 AAB24758	Plant SDF encoded
43	117.5	8.7	384	21 AAG42943	Arabidopsis thalia
44	117.5	8.7	384	21 AAG54911	Arabidopsis thalia
45	115	8.6	201	21 AAG38306	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAW74980	
ID	AAW74980 standard; Protein: 263 AA.
XX	
AC	AAW74980;
XX	
DT	25-JAN-1999 (first entry)
AC	
DE	Human secreted protein encoded by gene 106 clone HT3AM65.
XX	
KW	Human: secreted protein; testis; tumour; foetal brain tissue;
KW	fusion protein; cancer; central nervous system; seizure;
KW	diagnosis; neurodegenerative disease.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Misc-difference 263
FT	Location/Qualifiers
XX	
XX	W09839448-A2.
XX	
PD	11-SEP-1998.
XX	
PF	06-MAR-1998; 98NC-US04493.
XX	
PR	02-OCT-1997; 97US-0061060.
PR	07-MAR-1997; 97US-0038621.
PR	07-MAR-1997; 97US-0040161.
PR	07-MAR-1997; 97US-0040162.
PR	07-MAR-1997; 97US-0040163.
PR	07-MAR-1997; 97US-0040333.
PR	07-MAR-1997; 97US-0040334.

PR	07-MAR-1997	97US-0040336
PR	07-MAR-1997	97US-0040626
PR	11-APR-1997	97US-0043311
PR	11-APR-1997	97US-0043312
PR	11-APR-1997	97US-0043313
PR	11-APR-1997	97US-0043314
PR	11-APR-1997	97US-0043358
PR	11-APR-1997	97US-0043569
PR	11-APR-1997	97US-0043576
PR	11-APR-1997	97US-0043578
PR	11-APR-1997	97US-0043580
PR	11-APR-1997	97US-0043669
PR	11-APR-1997	97US-0043670
PR	11-APR-1997	97US-0043671
PR	11-APR-1997	97US-0043672
PR	11-APR-1997	97US-0043674
PR	23-MAY-1997	97US-0047582
PR	23-MAY-1997	97US-0047583
PR	23-MAY-1997	97US-0047584
PR	23-MAY-1997	97US-0047585
PR	23-MAY-1997	97US-0047586
PR	23-MAY-1997	97US-0047587
PR	23-MAY-1997	97US-0047588
PR	23-MAY-1997	97US-0047589
PR	23-MAY-1997	97US-0047590
PR	23-MAY-1997	97US-0047592
PR	23-MAY-1997	97US-0047593
PR	23-MAY-1997	97US-0047594
PR	23-MAY-1997	97US-0047595
PR	23-MAY-1997	97US-0047596
PR	23-MAY-1997	97US-0047597
PR	23-MAY-1997	97US-0047598
PR	23-MAY-1997	97US-0047599
PR	23-MAY-1997	97US-0047600
PR	23-MAY-1997	97US-0047601
PR	23-MAY-1997	97US-0047612
PR	23-MAY-1997	97US-0047613
PR	23-MAY-1997	97US-0047614
PR	23-MAY-1997	97US-0047615
PR	23-MAY-1997	97US-0047616
PR	23-MAY-1997	97US-0047618
PR	23-MAY-1997	97US-0047632
PR	23-MAY-1997	97US-0047633
PR	06-JUN-1997	97US-0048964
PR	06-JUN-1997	97US-0048974
PR	13-JUN-1997	97US-0049610
PR	16-JUL-1997	97US-0051926
PR	18-JUL-1997	97US-0052874
PR	18-AUG-1997	97US-0055724
PR	22-AUG-1997	97US-0055630
PR	22-AUG-1997	97US-0056331
PR	22-AUG-1997	97US-0056332
PR	22-AUG-1997	97US-0056336
PR	22-AUG-1997	97US-0056367
PR	22-AUG-1997	97US-0056637
PR	22-AUG-1997	97US-0056872
PR	22-AUG-1997	97US-0056873
PR	22-AUG-1997	97US-0056874
PR	22-AUG-1997	97US-0056875
PR	22-AUG-1997	97US-0056876
PR	22-AUG-1997	97US-0056877
PR	22-AUG-1997	97US-0056878
PR	22-AUG-1997	97US-0056879
PR	22-AUG-1997	97US-0056880

[illegible]

(HUMA-) HUMAN GENOME SCI INC.

Bednarik DF, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA, Feng P, Ferlie AM, Fischer CL, Florence KA, Greene JM, Hu JS, Ryan H, Laflair DW, Li Y, Moore PA, Ni J, Olsen HS, Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;

WPI; 1998-506364/43
N-PSDB; AAV59765.

New isolated human genes and the secreted polypeptide(s) they encode - useful for diagnosis and treatment of e.g. cancers, neurological disorders, immune diseases, inflammation or blood disorders

Claim 1; Page 686-687; 721pp; English.

This sequence represents a secreted human protein encoded by the nucleic acid molecule designated Gene 106 from the human cDNA clone H33AM65 (deposited as clone ATCC 97901 and ATCC 2090477). The gene can be used to generate fusion proteins by linking to the gene to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of the fused protein as compared to the human protein only. The invention relates to 186 novel genes and their fragments (nucleic acid sequences: AAV59511-V59612; amino acid sequences AAW74731-W75026) which are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Also, pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new polynucleotides. Specific uses are described for each of the 186 polynucleotides, based on which tissues they are most highly expressed in (see AAV59511 for described uses).

Sequence 263 AA;

Very Match	100.0%;	Score 1344;	DB 19;	Length 263;
	100.00	263	1	263-10

Matches	262;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	----	--------	----	------	----

1 MDCPALPPGWKKEEVIKRS

61 TGMKPSKLOKRNKRLNDPLNOKGKPDLTNTLPINQTSITKQPTKYTNTNPSNKVKS 120

Db 61 T G K M P S K L Q K N K Q R L R N D P L N Q N K G K P D L N T T L P I R Q T A S I F K Q P V T K V T N H P S N K V K S 120

121 DPQRMNEQPRQLFWERKLQGLSASDVTEQIKTMELPKGLQGVGPGSNDETLISAVASAL 180

Db 121 DPQRMNEQPRQLFWEKRLQGLSASDVTEQIKTMELPKGLQGVGPGSNDETLLSAVASAL 180

181 HTSSAPITGVSAVEKNPAWLNLSQPLCKAFIVTDEDIRKQEEERVQOVKKLEALMA 240

Db 181 HTSSAPITGQVSAVEKNPAVWLNTSQPLCKAFIVTDEDIRKQEEERVQQVKKLEALMA 240

OY 241 DILSRADTEEMDIEMDSGDEA 262
DB 241 DILSRADTEEMDIEMDSGDEA 262

RESULT 2

AAV73829
ID AAV73829 standard; Protein; 281 AA.
XX AAV73829;

XX 14-MAR-2000 (first entry)

XX Human prostate tumor EST fragment derived protein #16.

XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
XX treatment.

XX Homo sapiens.

XX DE19820190-A1.

XX 04-NOV-1999.

XX 28-APR-1998; 98DE-1020190.

XX 28-APR-1998; 98DE-1020190.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX WPI: 1999-621386/54.

XX DR N-PSDB: AA252863.

XX PT New human nucleic acid sequences from pancreatic tumors, and related
XX proteins

XX PS Claim 23; Page 315; 502pp; German.

CC This invention describes novel polypeptides and their encoding nucleic
CC acids derived from human pancreatic tumor tissue which have cytostatic
CC activity. The sequences are also useful in producing pharmaceutical
CC compositions for treatment of pancreatic tumors. AAV73814-Y74252
CC represent protein fragments encoded by the human pancreatic tumor cDNA
CC library derived expressed sequence tag (EST) sequences represented in
CC AA252858-253014.

XX Sequence 281 AA;

Query Match 100.0%; Score 1344; DB 20; Length 281;

Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;

Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDCPALPFGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNVTDLSSPFR 60
DB 20 MDCPALPFGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNVTDLSSPFR 79
OY 61 TGGKMPKSLQKKORLNDPLNOKNGKPDNTLPIROTASIFKOPVTKYTNHPSNKVKS 120
DB 80 TGGKMPKSLQKKORLNDPLNOKNGKPDNTLPIROTASIFKOPVTKYTNHPSNKVKS 139
OY 121 DPGRMNEQPRQLFWEKRLQGLSASDVTEQIITMELPKGLQGVGSGNDETLSSAVASAL 180
DB 140 DPGRMNEQPRQLFWEKRLQGLSASDVTEQIITMELPKGLQGVGSGNDETLSSAVASAL 199
OY 181 HTSSAPITGQVSAVEKNPAVWLNTSOPLCAKFIYDDEDIRKQERVOQVRKKLEBALMA 240
DB 200 HTSSAPITGQVSAVEKNPAVWLNTSOPLCAKFIYDDEDIRKQERVOQVRKKLEBALMA 259
OY 241 DILSRADTEEMDIEMDSGDEA 262
|||||

DB 260 DILSRADTEEMDIEMDSGDEA 281

RESULT 3

AAV48439
ID AAV48439 standard; Protein; 281 AA.

XX AAV48439;

XX 08-DEC-1999 (first entry)

XX Human prostate cancer-associated protein 136.

XX Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;
XX cancer; tissue specificity; human.

XX Homo sapiens.

XX DE19811194-A1.

XX 16-SEP-1999.

XX 10-MAR-1998; 98DE-1011194.

XX 10-MAR-1998; 98DE-1011194.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

XX WPI: 1999-519629/44.

XX DR N-PSDB: AA233533.

XX PT New nucleic acid sequences at high level in normal prostatic tissue and
XX encoded polypeptides, used to treat cancer and screen for therapeutic
XX agents

XX PS Claim 22; 176; 194pp; German.

CC This invention describes novel nucleic acid sequences (A) that are
CC expressed at high level in normal prostatic tissue. Polypeptides (I)
CC encoded by (A) are used: (a) for identifying agents for treatment of
CC prostatic cancer and (b) for therapy of prostate cancer, optionally
CC where expressed by gene therapy methods. (A) is also used to isolate
CC full-length genes (for gene therapy) and for recombinant production of
CC (I), which can be used to raise specific antibodies. (A) are identified
CC by assembly of ESTs (expressed sequence tags) before these are analyzed
CC for expression pattern (tissue specificity). This approach eliminates
CC many of the false results, as regards tissue specificity, associated
CC with known methods that use single (usually short) ESTs. AAV48304-Y48456
CC represent peptides encoded by the expressed sequence tags described in
CC the method of the invention.

XX Sequence 281 AA;

Query Match 100.0%; Score 1344; DB 20; Length 281;

Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;

Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDCPALPFGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNVTDLSSPFR 60
DB 20 MDCPALPFGMKKEEYIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNVTDLSSPFR 79
OY 61 TGGKMPKSLQKKORLNDPLNOKNGKPDNTLPIROTASIFKOPVTKYTNHPSNKVKS 120
DB 80 TGGKMPKSLQKKORLNDPLNOKNGKPDNTLPIROTASIFKOPVTKYTNHPSNKVKS 139
OY 121 DPGRMNEQPRQLFWEKRLQGLSASDVTEQIITMELPKGLQGVGSGNDETLSSAVASAL 180
DB 140 DPGRMNEQPRQLFWEKRLQGLSASDVTEQIITMELPKGLQGVGSGNDETLSSAVASAL 199
OY 181 HTSSAPITGQVSAVEKNPAVWLNTSOPLCAKFIYDDEDIRKQERVOQVRKKLEBALMA 240
DB 200 HTSSAPITGQVSAVEKNPAVWLNTSOPLCAKFIYDDEDIRKQERVOQVRKKLEBALMA 259
OY 241 DILSRADTEEMDIEMDSGDEA 262
|||||

Db 200 HTSSAPITGVSAAVEKNPAWMLNTSOLCKAFIVTDEDIRKQERVOYQVKKLEALMA 259
 QY 241 DILSRADTEMDIEMDSGDEA 262
 |||||||
 Db 260 DILSRADTEMDIEMDSGDEA 281

RESULT 4

AA14197 standard; Protein: 411 AA.

AA14197:

28-JUL-1999 (first entry)

Human DNA demethylase, dmtasel, protien sequence.

XX DNA demethylase; dmtasel; dmtase2; human; cancer cell; beta-thalassemia;
 KW DNA methylation pattern; methylation inhibitor; gene therapy; diagnosis;
 KM genetic defect correction; sickle cell anaemia.
 XX
 OS Homo sapiens.

PN W09924583-A1.

PD 20-MAY-1999.

PF 12-NOV-1998; 98WO-CA01059.

PR 11-MAY-1998; 98CA-2230991.

PR 12-NOV-1997; 97CA-2220805.

XX (UYMC-) UNIV MCGILL.

PI Bhattacharya S, Ramchandani S, Szyf M,

DR WPI; 1999-347283/29.

DR N-PSDB; AAX61218.

XX Human and murine DNA demethylases useful in the diagnosis of cancer

PT Disclosure; Fig 9c; 114pp; English.

XX This sequence is the human DNA demethylase, designated dmtasel, of
 CC the invention. The DNA demethylase is overexpressed in cancer cells.
 CC Expression of the demethylase cDNA is useful to alter DNA methylation
 CC patterns of DNA in vitro in cells or in vivo in humans, animals and
 CC plants. The cDNA is in antisense orientation to inhibit demethylase in
 CC cancer cells for therapeutic purposes. The demethylase is used to alter
 CC the differentiation state and to generate stem cells for therapeutics,
 CC cells for animal cloning and to improve expression of foreign genes. The
 CC cDNA can also be used for recombinant production of large amounts of the
 CC demethylase. The protein can be used to raise antibodies against
 CC demethylase. It can also be used for high throughput screening of
 CC demethylase inhibitors, and for obtaining the x-ray crystal structure.
 CC The demethylase cDNA and protein are also useful for changing the state
 CC of differentiation of a cell to allow gene therapy, stem cell selection
 CC or cell cloning; or for inhibiting methylation in cancer cells using
 CC vector mediated gene therapy. Antagonists or inhibitors of the
 CC demethylase can be used to manufacture medicaments for cancer treatment,
 CC for restoring an aberrant methylation patterns or changing methylation
 CC patterns in patient DNA. Change of the methylation pattern activates a
 CC silent gene permitting the correction of a genetic defect, such as a
 CC beta-thalassemia or sickle cell anaemia. The cDNA can be used as a
 CC template to design antisense oligonucleotides and ribozymes. The cDNA can
 CC also be used in two-hybrid systems in yeast to identify proteins
 CC interacting with demethylase. Determining the level of DNA methylase
 CC expression in a cell can be used as a method for diagnosis of cancer,
 CC where overexpression is indicative of cancer cells.

XX Sequence 411 AA;

Query Match 100.0%; Score 1344; DB 20; Length 411;

Best Local Similarity 100.0%; Pred. No. 2.8e-119;
 Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCPALPGMKKEVIRKSGLSAGSDVYVYFSPGKKFRSKQRLARYIGNTVDLSFDFR 60
 |||||||
 Db 150 MDCPALPGMKKEVIRKSGLSAGSDVYVYFSPGKKFRSKQRLARYIGNTVDLSFDFR 209
 QY 61 TGMKMSKLOKNQRLRNPLNKNKGPDLNLTLPFRORASIFKOPVTVTHNPNKYS 120
 |||||||
 Db 210 TGMKMSKLOKNQRLRNPLNKNKGPDLNLTLPFRORASIFKOPVTVTHNPNKYS 269
 QY 121 DPGRMNEOPROLFEWERLGLSADVTEQIIKTMELPKGLGQVGSNDETLLSAVASAL 180
 |||||||
 Db 270 DPGRMNEOPROLFEWERLGLSADVTEQIIKTMELPKGLGQVGSNDETLLSAVASAL 329
 QY 181 HTSSAPITGVSAAVEKNPAWMLNTSOLCKAFIVTDEDIRKQERVOYQVKKLEALMA 240
 |||||||
 Db 330 HTSSAPITGVSAAVEKNPAWMLNTSOLCKAFIVTDEDIRKQERVOYQVKKLEALMA 389
 QY 241 DILSRADTEMDIEMDSGDEA 262
 |||||||
 Db 390 DILSRADTEMDIEMDSGDEA 411

RESULT 5

AAB99915 standard; Protein: 411 AA.

AC AAB99915;

DT 26-SEP-2001 (first entry)

XX Human protein sequence SEQ ID NO:1.

XX Differentiation; heart muscle cell; cytokine; transcription factor;
 KW proliferation; surface antigen; heart disease; cardiomyocyte;
 KW bone marrow; umbilical blood cell; heart muscle degeneration;
 KW myocardial infarction.

OS Homo sapiens.

PN W0200148150-A1.

PD 05-JUL-2001.

PF 02-NOV-2000; 2000WO-JP07741.

PR 28-DEC-1999; 99JP-0372826.

PR 28-FEB-2000; 2000WO-JP01148.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PI Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

DR WPI; 2001-425655/45.

DR N-PSDB; AAH44351.

XX Cells capable of differentiating into cardiomyocytes and originating in
 PT bone marrow or umbilical blood cells for study of cardiomyocyte
 PT differentiation and treatment of heart disease
 XX Claim 22; Page 84-86; 187pp; Japanese.

XX The present invention describes cells originating in bone marrow or
 CC umbilical blood cells which are capable of differentiating into
 CC cardiomyocytes. Also described are: (1) cardiomyocytes produced by the
 CC differentiation of the cells; (2) a method for carrying out the
 CC differentiation into cardiomyocytes, regulated by a promotional and/or
 CC inhibitory factor; (3) a method for the differentiation of the cells
 CC into cell types other than cardiomyocytes; (4) drug compositions
 CC promoting the formation of heart muscle and regeneration of heart tissue
 CC which contain the cells; (5) a method for the production of antibodies

CC which recognise the cells, especially antibodies which recognise a
CC surface antigen on the cells; (6) a method for screening factors which
CC promote the proliferation of the cells; (7) a method for immortalising
CC the cells by expressing telomerase in them; (8) drug compositions for
CC the treatment of heart disease which contain the immortalised cells; and
CC (9) cell-free supernatant from the culture of the immortalised cells; and
CC promoting their differentiation into cardiomyocytes. The cells are used
CC in the treatment of diseases involving heart muscle degeneration, such
CC as myocardial infarction and in the study of cardiomyocyte
CC differentiation. AAH44351 to AAH44409 and AAB99915 to AAB99935 represent
CC sequences used in the exemplification of the present invention.
XX

Sequence 411 AA:

Query Match 100.0%; Score 1344; DB 22; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-119;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCPALPPGKKKEEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDSLSPDFR 60
DB 150 MDCPALPPGKKKEEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDSLSPDFR 209
QY 61 TGGKMPKSLQKKORLNDPLNKKGPDLNTTLPPIROTASIFKOPVTKYTNHPSNKVKS 120
DB 210 TGGKMPKSLQKKORLNDPLNKKGPDLNTTLPPIROTASIFKOPVTKYTNHPSNKVKS 269
QY 121 DPORMNEQPOLFEWKRLOGLSASDVTEQIITMELPKGLGVGPGSNDETLLSAVASAL 180
DB 270 DPORMNEQPOLFEWKRLOGLSASDVTEQIITMELPKGLGVGPGSNDETLLSAVASAL 329
QY 181 HTSSAPITGOVSAAVEKNPAVWLNTSOPLCRAFIYDDEDIRKQEBRVQVQRKLEALMA 240
DB 330 HTSSAPITGOVSAAVEKNPAVWLNTSOPLCRAFIYDDEDIRKQEBRVQVQRKLEALMA 369
QY 241 DILSRADTEEMDIEMDSGEA 262
DB 390 DILSRADTEEMDIEMDSGEA 411

RESULT 6
AAG64314

ID AAG64314 standard; Protein; 411 AA.

XX AAG64314;

DT 24-SEP-2001 (first entry)

XX Human protein #1.

DE Angiogenesis; cardiact; cell differentiating agent; bone marrow;

KW heart muscle cell; heart disease; human.

XX Homo sapiens.

OS MO200148149-A1.

PN 05-JUL-2001.

PD 28-FEB-2000; 2000WO-JP01148.

PF 28-DEC-1999; 99JP-0372826.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PA Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K;

PI WPI: 2001-418252/44.

DR N-PSDB; AAH49386.

XX New adult bone marrow-originated cells capable of differentiating into
PT heart muscle cells, applicable as remedies for various heart diseases
XX particularly with damaged heart muscle accompanying degeneration -

PS Claim 12; Pages 55-57; 158pp; Japanese.

XX The present invention relates to cells isolated from bone marrow, which
CC are capable of at least differentiating into heart muscle cells. The
CC cells are applicable as remedies for various heart diseases particularly
CC with damaged heart muscle accompanying degeneration. The present sequence
XX was used to illustrate the present invention.

Sequence 411 AA:

Query Match 100.0%; Score 1344; DB 22; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-119;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCPALPPGKKKEEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDSLSPDFR 60
DB 150 MDCPALPPGKKKEEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNTVDSLSPDFR 209
QY 61 TGGKMPKSLQKKORLNDPLNKKGPDLNTTLPPIROTASIFKOPVTKYTNHPSNKVKS 120
DB 210 TGGKMPKSLQKKORLNDPLNKKGPDLNTTLPPIROTASIFKOPVTKYTNHPSNKVKS 269
QY 121 DPORMNEQPOLFEWKRLOGLSASDVTEQIITMELPKGLGVGPGSNDETLLSAVASAL 180
DB 270 DPORMNEQPOLFEWKRLOGLSASDVTEQIITMELPKGLGVGPGSNDETLLSAVASAL 329
QY 181 HTSSAPITGOVSAAVEKNPAVWLNTSOPLCRAFIYDDEDIRKQEBRVQVQRKLEALMA 240
DB 330 HTSSAPITGOVSAAVEKNPAVWLNTSOPLCRAFIYDDEDIRKQEBRVQVQRKLEALMA 369
QY 241 DILSRADTEEMDIEMDSGEA 262
DB 390 DILSRADTEEMDIEMDSGEA 411

RESULT 7
AAG64844

ID AAG64844 standard; Protein; 411 AA.

XX AAG64844;

DT 21-SEP-2001 (first entry)

XX Heart muscle cell differentiation related protein SEQ ID NO: 1.

DE Heart muscle cell; human; cell differentiation; heart disease.

KW Homo sapiens.

OS MO200148151-A1.

PN 05-JUL-2001.

PD 27-DEC-2000; 2000WO-JP09323.

PF 28-DEC-1999; 99JP-0372826.

PR 28-FEB-2000; 2000WO-JP01148.

XX 02-NOV-2000; 2000WO-JP07741.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PA Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

PI Yamada Y;

DR WPI: 2001-425656/45.

XX N-PSDB; AAH48220.

XX Cells capable of differentiating into cardiomyocytes and originating in
PT bone marrow or umbilical blood cells for study of cardiomyocyte
XX differentiation and treatment of heart disease -
PS Claim 28; Page 90-92; 183pp; Japanese.

CC The present invention provides cells originating in the human bone marrow
 CC or umbilical blood cells which are capable of differentiating into
 CC cardiomyocytes. These cells are useful in the treatment of diseases
 CC involving heart muscle degeneration, such as myocardial infarction, and
 CC the study of cardiomyocyte differentiation. The present sequence is
 CC a protein described in the exemplification of the invention.

SO Sequence 411 AA;

Query Match 100.0%; Score 1344; DB 22; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2,8e-119;
 Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDCPALPGWKKEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNVDSLSPDR 60
 DB 150 MDCPALPGWKKEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNVDSLSPDR 209
 QY 61 TGMKMPKLOKRNKORLNDPLNOKNGKRPDLNTLPIRQTASIFKQPYTKVTNHPNKKVKS 120
 DB 210 TGMKMPKLOKRNKORLNDPLNOKNGKRPDLNTLPIRQTASIFKQPYTKVTNHPNKKVKS 269
 QY 121 DPORMNEOPROLFWEKRLQGLSASDVTEQIIKTMELPKLGQVGPSSNDETLISAVAL 180
 DB 270 DPORMNEOPROLFWEKRLQGLSASDVTEQIIKTMELPKLGQVGPSSNDETLISAVAL 329
 QY 181 HTSSAPITGCVSAVEKNPAWMLNTSOPLCARIVTDEDIRKOEERVOQVRKLEALMA 240
 DB 330 HTSSAPITGCVSAVEKNPAWMLNTSOPLCARIVTDEDIRKOEERVOQVRKLEALMA 389
 QY 241 DILSRAADTEEMDIEMDSGDEA 262
 DB 390 DILSRAADTEEMDIEMDSGDEA 411

RESULT 8
 AAY14199
 ID AAY14199 standard; Protein: 414 AA.

XX AAY14199;

DT 28-JUL-1999 (first entry)

DE Mouse DNA demethylase, dmtase1, protein sequence.

KW DNA demethylase; dmtase1; mouse; cancer cell; beta-thalassemia;
 KW DNA methylation pattern; methylation inhibitor; gene therapy; diagnosis;
 KW genetic defect correction; sickle cell anaemia.

OS Mus sp.

PN WO9924583-A1.

PD 20-MAY-1999.

PF 12-NOV-1998; 98WO-CA01059.

PR 11-MAY-1998; 98CA-2230991.

PR 12-NOV-1997; 97CA-2220805.

XX (UTWC-) UNIV MCGILL.

PI Bhattacharya S, Ramchandani S, Szyf M;

DR WPI; 1999-347283/29.

XX N-PSDB; AAX61220.

PT Human and murine DNA demethylases useful in the diagnosis of cancer

PS Disclosure; Fig 9k; 114pp; English.

CC This sequence is the mouse DNA demethylase, designated dmtase1, of
 CC the invention. The DNA demethylase is overexpressed in cancer cells.
 CC Expression of the demethylase cDNA is useful to alter DNA methylation

CC patterns of DNA in vitro in cells or in vivo in humans, animals and
 CC plants. The cDNA is in antisense orientation to inhibit demethylase in
 CC cancer cells for therapeutic purposes. The demethylase is used to alter
 CC the differentiation state and to generate stem cells for therapeutics.
 CC cells for animal cloning and to improve expression of foreign genes. The
 CC cDNA can also be used for recombinant production of large amounts of the
 CC demethylase. The protein can be used to raise antibodies against
 CC demethylase. It can also be used for high throughput screening of
 CC demethylase inhibitors, and for obtaining the x-ray crystal structure.
 CC The demethylase cDNA and protein are also useful for changing the state
 CC of differentiation of a cell to allow gene therapy, stem cell selection
 CC or cell cloning; or for inhibiting methylation in cancer cells using
 CC vector mediated gene therapy. Antagonists or inhibitors of the
 CC demethylase can be used to manufacture medicaments for cancer treatment,
 CC for restoring an aberrant methylation patterns or changing methylation
 CC patterns in patient DNA. Change of the methylation pattern activates a
 CC silent gene permitting the correction of a genetic defect, such as
 CC beta-thalassemia or sickle cell anaemia. The cDNA can be used as a
 CC template to design antisense oligonucleotides and ribozymes. The cDNA can
 CC also be used in two-hybrid systems in yeast to identify proteins
 CC interacting with demethylase. Determining the level of DNA methylase
 CC expression in a cell can be used as a method for diagnosis of cancer,
 CC where overexpression is indicative of cancer cells.

SO Sequence 414 AA;

Query Match 98.7%; Score 1327; DB 20; Length 414;
 Best Local Similarity 98.5%; Pred. No. 1.2e-117;
 Matches 258; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDCPALPGWKKEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNVDSLSPDR 60
 DB 153 MDCPALPGWKKEVIRKSGLSAGKSDVYFSPSGKKFRSKPOLARYLGNVDSLSPDR 212

QY 61 TGMKMPKLOKRNKORLNDPLNOKNGKRPDLNTLPIRQTASIFKQPYTKVTNHPNKKVKS 120
 DB 213 TGMKMPKLOKRNKORLNDPLNOKNGKRPDLNTLPIRQTASIFKQPYTKVTNHPNKKVKS 272

QY 121 DPORMNEOPROLFWEKRLQGLSASDVTEQIIKTMELPKLGQVGPSSNDETLISAVAL 180
 DB 273 DPORMNEOPROLFWEKRLQGLSASDVTEQIIKTMELPKLGQVGPSSNDETLISAVAL 332

QY 181 HTSSAPITGCVSAVEKNPAWMLNTSOPLCARIVTDEDIRKOEERVOQVRKLEALMA 240
 DB 333 HTSSAPITGCVSAVEKNPAWMLNTSOPLCARIVTDEDIRKOEERVOQVRKLEALMA 392

QY 241 DILSRAADTEEMDIEMDSGDEA 262
 DB 393 DILSRAADTEEMDIEMDSGDEA 414

RESULT 9
 AAE22578
 ID AAE22578 standard; Protein: 282 AA.

XX AAE22578;

DT 26-JUL-2002 (first entry)

DE Xenopus laevis MBD3 protein.

KW Gene expression; cellular chromatin; methyl CpG binding domain; cancer;
 KW localisation domain; diabetic retinopathy; ischaemia; HIV infection;
 KW human immuno deficiency virus; macular degeneration; vascular disease;
 KW rheumatoid arthritis; psoriasis; Alzheimer's disease; muscular dystrophy;
 KW sickle cell anaemia; stroke; neurodegenerative disease; cystic fibrosis;
 KW gene therapy; cytototoxic; antidiabetic; ophthalmological; vasotrophic;
 KW neuroprotective; nootropic; cerebroprotective; antibacterial; antifungal;
 KW antiviral; MBD3 protein.

OS Xenopus laevis.

XX Key Location/Qualifiers
 FH

QY	122	POGMNQPPQOLFEWERKLOGLASDVTEOIIKTMELPKLOGVGSGNDEFTLLSAVASALH	181
Db	126	POAAVQPPQOLFEWERKLSGLNFDIAEELVVTMELPKLOGVGSGCTDEFTLLSAISALH	185
QY	182	TSSAPITGVGVAASAEKRNPAVMTNSQPLCKAFIYTDDEIRKQEEHVOQVKKLTLEALMAD	241
Db	186	TSMPITGSLAASAEKRNPAVMTNSQPLCKAFIYTDDEIRKQEEHVOQVKKLTLEALMAD	245
QY	242	ILSRADTEE-----MDIEMSGDE 261	
Db	246	MLAHVEISKDGAPLDRKIDDEE 270	
RESULT 10			
AAE22577	ID	AAE22577 standard; Protein; 303 AA.	
XX	AAE22577;		
AC	XX		
XX	DT	26-JUL-2002 (first entry)	
XX	DE	Xenopus laevis MBD3 LF protein.	
XX	XX	Gene expression; cellular chromatin; methyl CpG binding domain; cancer; localization domain; diabetic retinopathy; ischemia; HIV infection; human immuno deficiency virus; macular degeneration; vascular disease; rheumatoid arthritis; psoriasis; Alzheimer's disease; muscular dystrophy; sickle cell anaemia; stroke; neurodegenerative disease; cystic fibrosis; gene therapy; cytosolic; antidiabetic; ophthalmological; vasotropic; neuroprotective; nontropic; cerebroprotective; antibacterial; antifungal; antiviral; MBD3 protein.	
XX	OS	Xenopus laevis.	
XX	XX		
FH	Key	Location/Qualifiers	
FT	Binding-site	1..90	
FT		/label= Methyl_CpG_binding_domain	
FT	Region	7..9	
FT		/label= Beta1_helix	
FT	Region	15..21	
FT		/label= Beta2_helix	
FT	Region	22..31	
FT		/label= L1_loop	
FT	Region	32..37	
FT		/label= Beta3_helix	
FT	Region	61..63	
FT		/label= Beta4_helix	
FT	Region	67..73	
FT		/label= Alpha1_helix	
FT	Region	74..80	
FT		/label= L2_loop	
FT	Region	82..89	
FT		/label= Hairpin_loop	
PN	MO200226960-A2.		
XX	XX		
XX	PD	04-APR-2002.	
XX	PF	28-SEP-2001; 2001MO-US42377.	
XX	PR	29-SEP-2000; 2000US-236884P.	
XX	PA	(SANG-) SANGAMO BIOSCIENCES INC.	
XX	PI	Wolffe AP, urnov F, Lai A, Raschke E;	
XX	XX	WPI; 2002-372124/40.	
XX	XX		
PT	XX	Compartentalizing a region of interest in cellular chromatin which facilitates the modulation of the expression of a gene comprises contacting the gene with a composition comprising a localized agent-	

AC AAY07107;
 XX 02-JUL-1999 (first entry)
 XX
 DE Colon cancer associated antigen precursor sequence.
 XX
 KW Cancer associated antigen; diagnosis; research; treatment; human;
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 KW prostate cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9904265-A2.
 XX
 PD 28-JAN-1999.
 XX
 PF 15-JUL-1998; 98WO-US14679.
 XX
 PR 22-JUN-1998; 98US-0102322.
 PR 17-JUL-1997; 97US-0896164.
 PR 10-OCT-1997; 97US-0061599.
 PR 10-OCT-1997; 97US-0061765.
 PR 10-OCT-1997; 97US-0948705.
 PR 11-OCT-1997; 97GB-0021697.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Chen Y, Gout I, Gure A, O'Hare M, Ohta Y, Old LJ;
 PI Pfeundschnuh M, Sahln U, Scanlan MJ, Stockert E;
 PI Tureci O;
 XX
 DR WPI: 1999-132448/11.
 XX
 PT New isolated cancer associated nucleic acids and polypeptides -
 PT isolated using sera from cancer patients, used to develop products
 PT for the diagnosis, monitoring or treatment of cancers
 XX
 PS Disclosure: Page 677; 787pp; English.
 XX
 CC The invention relates to a method for diagnosing a disorder characterised
 CC by expression of a human cancer associated antigen precursor coded for by
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
 CC biological sample isolated from a subject with an agent that specifically
 CC binds to the NAM, an expression product or a fragment of an expression
 CC product complexed with an HLA molecule; and (b) determining the
 CC interaction between the agent and the NAM or the expression product as a
 CC determination of the disorder. The products and methods can be used in
 CC the diagnosis, monitoring, research, or treatment of conditions
 CC characterised by the expression of various cancer associated antigens.
 CC The invention provides nucleic acid sequences and encoded polypeptides
 CC which are cancer associated antigen precursors expressed in human breast
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC lung cancer.
 XX
 SQ Sequence 200 AA;
 Query Match 75.4%; Score 1013; DB 20; Length 200;
 Best Local Similarity 100.0%; Pred. No. 3.2e-88;
 Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 KMPFSKQKKKQRLNDPLNOKNGKPDLTPTLPIROTASIFKQPVYKVNHPNSNKVKSOP 122
 DB 1 KMPFSKQKKKQRLNDPLNOKNGKPDLTPTLPIROTASIFKQPVYKVNHPNSNKVKSOP 60
 QY 123 ORNNEOPROLFWFKRLQGLSASDVTEQIIKTMLPGLGVGPGSNDFTLLSAVASALHT 182
 DB 61 ORNNEOPROLFWFKRLQGLSASDVTEQIIKTMLPGLGVGPGSNDFTLLSAVASALHT 120
 QY 183 SSAPITGOVSAAEKNPAWLNTSOPLCFAFYTDDEDIKQBERVOOVYKKLEELMDI 242
 DB 121 SSAPITGOVSAAEKNPAWLNTSOPLCFAFYTDDEDIKQBERVOOVYKKLEELMDI 180
 QY 243 LSRADTEEMDIDMSGDGA 262

DB 181 LSRADTEEMDIDMSGDGA 200
 RESULT 13
 AAY14200
 ID AAY14200 standard; Protein; 285 AA.
 XX
 AC AAY14200;
 XX
 XX 28-JUL-1999 (first entry)
 DE Mouse DNA demethylase, dMTase2, protein sequence.
 KW DNA demethylase; dMTase2; mouse; cancer cell; beta-thalassemia;
 KW DNA methylation pattern; methylation inhibitor; gene therapy; diagnosis;
 KW genetic defect correction; sickle cell anaemia.
 XX
 OS Mus sp.
 XX
 PN WO9924583-A1.
 XX
 PD 20-MAY-1999.
 XX
 PF 12-NOV-1998; 98WO-CA01059.
 XX
 PR 11-MAY-1998; 98CA-2230991.
 PR 12-NOV-1997; 97CA-2220805.
 XX
 PA (UYWC-) UNIV MCGILL.
 XX
 PI Bhattacharya S, Ramchandani S, Szyf M;
 PI WPI: 1999-347283/29.
 DR N-PSDB; AAX61221.
 XX
 PT Human and murine DNA demethylases useful in the diagnosis of cancer
 PT
 XX
 PS Disclosure: Fig 9n; 114pp; English.
 XX
 CC This sequence is the mouse DNA demethylase, designated dMTase2, of
 CC the invention. The DNA demethylase is overexpressed in cancer cells.
 CC Expression of the demethylase cDNA is useful to alter DNA methylation
 CC patterns of DNA in vitro in cells or in vivo in humans, animals and
 CC plants. The cDNA is in antisense orientation to inhibit demethylase in
 CC cancer cells for therapeutic purposes. The demethylase is used to alter
 CC the differentiation state and to generate stem cells for therapeutics,
 CC cells for animal cloning and to improve expression of foreign genes. The
 CC cDNA can also be used for recombinant production of large amounts of the
 CC demethylase. The protein can be used to raise antibodies against
 CC demethylase. It can also be used for high throughput screening of
 CC demethylase inhibitors, and for obtaining the x-ray crystal structure.
 CC The demethylase cDNA and protein are also useful for changing the state
 CC of differentiation of a cell to allow gene therapy, stem cell selection
 CC or cell cloning; or for inhibiting methylation in cancer cells using
 CC vector mediated gene therapy. Antagonists or inhibitors of the
 CC demethylase can be used to manufacture medicaments for cancer treatment,
 CC for restoring an aberrant methylation patterns or changing methylation
 CC patterns in patient DNA. Change of the methylation pattern activates a
 CC silent gene permitting the correction of a genetic defect, such as
 CC beta-thalassemia or sickle cell anaemia. The cDNA can be used as a
 CC template to design antisense oligonucleotides and ribozymes. The cDNA can
 CC also be used in two-hybrid systems in yeast to identify proteins
 CC interacting with demethylase. Determining the level of DNA methylation
 CC expression in a cell can be used as a method for diagnosis of cancer,
 CC where overexpression is indicative of cancer cells.
 XX
 SQ Sequence 285 AA;
 Query Match 75.3%; Score 1012.5; DB 20; Length 285;
 Best Local Similarity 75.3%; Pred. No. 6e-88;
 Matches 189; Conservative 35; Mismatches 24; Indels 3; Gaps 1;

xx The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any special methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

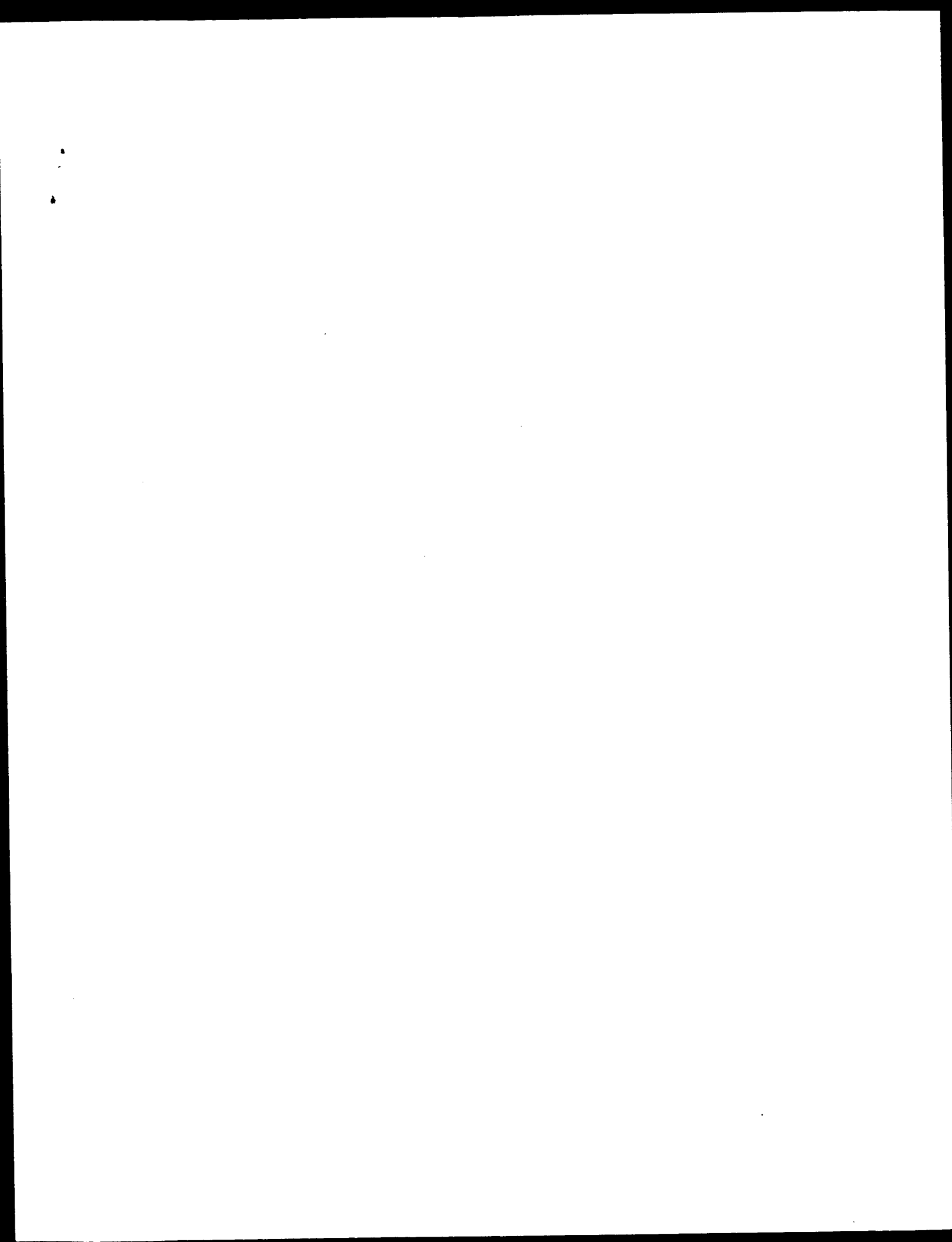
xx
 S0 Sequence 218 AA;

Query Match 55.0%; Score 739; DB 22; Length 218;
 Best Local Similarity 70.7%; Pred. No. 4,4e-62;

Matches 145; Conservative 24; Mismatches 24; Indels 12; Gaps 2;

Qy 69 LQKKQRRLNDPLNQNGKGPDLNTTLPTRQTASIFKQPVTKVTNHPNKNKVSDDPQRMNEQ 128
 Db 1 MNKSRQVRKRDSSNQVKGKDPDLNTALPVRQTASIFKQPVTKVTNHPNKNKVSDDPQKAYDQ 60
 Qy 129 PROLFWEKRLQGLSASDVTOITMELPKGLGVGPGSNDFTLLSAVASALHTSCAPIT 188
 Db 61 PROLFWEKRLQGLSASDVTOITMELPKGLGVGPGSNDFTLLSAVASALHTSCAPIT 120
 Qy 189 GQVSAAVEKNPAAVWLNTPSPLCAFIYTDIDIRKQERVOQVRRKLEALMADILSR--- 245
 Db 121 GQLSAAVEKNPAAVWLNTPSPLCAFIYTDIDIRKQERVOQVRRKLEALMADILSR--- 180
 Qy 246 -AADTE-----ENDIEMDSGDE 261
 Db 181 LARDGEAPLDKACAEDEDEDEEE 205

Search completed: March 12, 2003, 05:40:35
 Job time : 39.1516 secs




```

CC -----
DR EMBL: L37298; AAC32737.1; -
DR EMBL: Y12643; CAA73190.1; -
DR EMBL: X96886; CAA68001.1; -
DR EMBL: AF030876; AAC08757.1; -
DR EMBL: AF031078; AAC08758.1; -
DR EMBL: AJ132917; CAA64446.1; -
DR EMBL: X89430; CAA61599.1; -
DR EMBL: X94628; CAA64331.1; -
DR TRANSFAC: T04936; -
DR Genew: HGNC:6990; MECP2.
DR MIM: 300005; -
DR MIM: 312750; -
DR InterPro: IPR000637; AT-hook.
DR InterPro: IPR001739; Methyl-CpG_bind.
DR Pfam: PF01429; MBD; 1.
DR SMART: SM00384; AT_hook; 1.
DR SMART: SM00391; MBD; 1.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;
KW Disease mutation; Polymorphism.
FT DOMAIN 96 149 MBD.
FT DOMAIN 277 283 POLY-ALA.
FT DOMAIN 366 372 POLY-HIS.
FT DOMAIN 384 393 POLY-PRO.
FT VARIANT 101 101 P -> R (IN RTT).
FT VARIANT 106 106 /FTID-VAR_010276.
FT VARIANT 106 106 R -> W (IN RTT).
FT VARIANT 124 124 /FTID-VAR_010272.
FT VARIANT 124 124 L -> F (IN RTT).
FT VARIANT 133 133 /FTID-VAR_010277.
FT VARIANT 133 133 R -> C (IN RTT).
FT VARIANT 134 134 /FTID-VAR_010273.
FT VARIANT 134 134 S -> C (IN RTT).
FT VARIANT 140 140 /FTID-VAR_010278.
FT VARIANT 140 140 A -> V (IN RTT; AFFECTS WOMEN AND MEN).
FT VARIANT 152 152 /FTID-VAR_010279.
FT VARIANT 152 152 P -> R (IN RTT).
FT VARIANT 155 155 /FTID-VAR_010280.
FT VARIANT 155 155 F -> S (IN RTT).
FT VARIANT 158 158 /FTID-VAR_010274.
FT VARIANT 201 201 T -> M (IN RTT).
FT VARIANT 201 201 A -> V.
FT VARIANT 201 201 /FTID-VAR_010275.
FT VARIANT 306 306 /FTID-VAR_010281.
FT VARIANT 306 306 R -> C (IN RTT).
FT VARIANT 397 397 /FTID-VAR_010282.
FT VARIANT 397 397 E -> K.
FT CONFLICT 72 75 /FTID-VAR_010283.
FT CONFLICT 290 290 PAVP -> RLC (IN REF. 6).
FT CONFLICT 290 290 E -> G (IN REF. 3).
SQ SEQUENCE 486 AA; 52440 MW; E6A3333AEDA566 CRC64;
Query Match 9.5%; Score 205.5; DB 1; Length 486;
Best Local Similarity 28.5%; Pred. No. 2.4e-05;
Matches 85; Conservative 38; Mismatches 104; Indels 71; Gaps 15;
QY 117 GGGGAPRRVPPPSGAGCPGPRATESGGRMDCPALPPGKKKEVIRKSGLSAGKSD 176
DB 67 GSGSAP-----AVPEASASPKQRSLIRDRGPMYDDPTLPEGWTRKLKOKKSGRSAGKTD 121
QY 177 VYFSPSGKKFRSKPOLARY---LGMT-VDSLSEDFR-TGKMPKSLQKKQKQLRNDP-- 229
DB 122 VYLINQGKAFKSKVELIAFEKVGDITSDPNDFTVYGRGSPSRRECKPPKPKPSPA 181
QY 230 -LNQNGKP-DLNTTLPPIKOTASIFKQPYTKVTNHPNKNVKSDDPORMNDOPQLFWKMR 286
DB 182 PGTGRGRGRKSGGTTRPKATSEGVQ--VKRYLE-----KSPGLLIKMPQIJSFGCK 233
QY 287 LQGLSADYTE-QIKR-----TWELPKGLQGVGPGSNDFTLLSAVASALHTSSA 334
DB 234 AEGGGAITTSQVAVIKRPGKRRKAADPOAIFK-KRGRKPG-----SVVAAA----- 279
QY 335 PITGQSAVAEKNPAAVWLNTSQPLCAFIYTDIDIKQGERVQGVAK-KLEELIADTI 391

```

```

DB 280 -----AAEAKKKA-----VKESIRSVQETVLPINKRKRTREVSIEV 316
RESULT 7
GRPL_ORYSA
ID GRPL_ORYSA STANDARD; PRT; 165 AA.
AC P25074;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Glycine-rich cell wall structural protein 1 precursor.
GN GRP-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Indica-IR36;
RA MEDLINE=91370862; PubMed=1716496;
RX Lei M., Wu R.;
RT "A novel glycine-rich cell wall protein gene in rice.";
RL Plant Mol. Biol. 16:187-198(1991).
CC -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Cell wall (potential).
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53596; CAA37665.1; -
DR PIR: S13385; KNRG1.
KW Cell wall; Structural protein; Repeat; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 165 GLYCINE-RICH CELL WALL STRUCTURAL
FT DOMAIN 31 159 PROTEIN 1.
FT REPEAT 56 62 GLY-RICH.
FT REPEAT 93 99 R2 (TYR-RICH).
FT REPEAT 132 138 R2 (TYR-RICH).
SQ SEQUENCE 165 AA; 13536 MW; E36CE31C3650AC9A CRC64;
Query Match 9.3%; Score 200.5; DB 1; Length 165;
Best Local Similarity 40.8%; Pred. No. 1.4e-05;
Matches 51; Conservative 6; Mismatches 59; Indels 9; Gaps 3;
QY 17 GSAAGSGAGDSDAIEGCGGSAAPSPVSGVRREGAR-GGGRGRGMKQAGRGVCG 75
DB 32 GSGSGGGGGGGGGGGG---GNGSGSGSGGYGKAAGGSGGGGGG 83
QY 76 RGRGRGRGRGRGRGRGRGRPPSGSGGLGCGGGCGGGSGGGAAPRRVPPPSGAG 135
DB 84 GNGSGSGSGGYGYGQNGAAGGSGGGGGGGGGGGGGSGGSGGYGYKGGCGG 143
QY 136 PGPGR 140
DB 144 GGGGG 148
RESULT 8
LORI_MOUSE
ID LORI_MOUSE STANDARD; PRT; 481 AA.
AC P18165;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Loricrin.

```


ID	GRP_ARATH	STANDARD:	PRT:	338 AA.
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	01-AUG-1992 (Rel. 23, Last annotation update)			
DE	Glycine-rich cell wall structural protein precursor.			
OC	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eustoids II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=cv. Columbia;			
RX	MEDLINE=92003708; PubMed=1912511;			
RA	Outgley F., Villiot M.L., Maché R.;			
RT	"Nucleotide sequence and expression of a novel glycine-rich protein			
RT	gene from Arabidopsis thaliana."			
RL	Plant Mol. Biol. 17:949-952(1991).			
CC	-1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: Cell wall (Potential).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X58338; CAA41249.1; -			
DR	PIR; S17732; KMMO.			
KW	Cell wall; Structural protein; Repeat; Signal.			
FT	SIGNAL	1	20	POTENTIAL.
FT	CHAIN	21	338	GLYCINE-RICH CELL WALL STRUCTURAL
FT				PROTEIN.
FT	DOMAIN	21	338	GLY-RIICH.
SO	SEQUENCE	338 AA;	23891 MW;	046A6E8C1A4E89EB CRC64;
Query Match		8.8%;	Score 191.5;	DB 1; Length 338;
Best Local Similarity		40.9%;	Pred. No. 0.0001;	
Matches	61; Conservative	2; Mismatches	53; Indels	33; Gaps
OY	6 GGGRCCEPDEGESAGSGAGGDSALF-----QGCGSALAPSPVGYRREBARGGG 58			
Db	193 GGG-----AGCGAGGCGGAGCGGGLGCGHGGCGGAGGAGGLGCGAGGCT--GGGFGGG 243			
OY	59 RGRGRWKQAGNG--GGVCGRGRGRGRGRGRGRGRGRPRPSGGSGGLGDDGGCG--CG 114			
Db	244 AGGGAGGAGGAGGFGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG--GGHGGVGGGFGG 302			
OY	115 GSG--GGCAPRREPVPSPSGAGGPPRG 140			
Db	303 GSGGFGGGA-----GGAGGAGGAGG 322			
RESULT 14				
ID	RB87_DROME			
ID	RB87_DROME	STANDARD:	PRT:	386 AA.
AC	P48810.			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Heterogeneous nuclear ribonucleoprotein 87F (HRP36.1 protein) (P11			
DE	protein).			
GN	HRB87F OR HRP36.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-S; and Canton-S; TISSUE=Ovary;
 RX MEDLINE=91187645; PubMed=1849257;
 RA Haynes S.R., Johnson D., Raychaudhuri G., Beyer A.L.;
 RT "The Drosophila Hrb7F gene encodes a new member of the A and B hnRNP
 RL protein group.";
 RL Nucleic Acids Res. 19:25-31(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=92112968; PubMed=1730754;
 RA Matunis E.L., Matunis M.J., Dreyfuss G.;
 RT "Characterization of the major hnRNP proteins from Drosophila
 RL melanogaster.";
 RL J. Cell Biol. 116:257-269(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S; TISSUE=Embryo;
 RX MEDLINE=92020124; PubMed=1719937;
 RA Hovemann B.T., Dessen E., Mechler H., Mack E.;
 RT "Drosophila snRNP associated protein P11 which specifically binds to
 RL heat shock puff 93D reveals strong homology with hnRNP core protein
 RL A1.";
 RL Nucleic Acids Res. 19:4909-4914(1991).
 CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF RIBONUCLEOSOMES. COULD BE
 CC NEEDED TO ORGANIZE A CONCENTRATION GRADIENT OF A DORSALIZING
 CC MORPHOGEN (DM) ORIGINATING IN THE GERMINAL VESICLE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR AND/OR CYTOPLASMIC.
 CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X54803; CAA38574.1; -
 DR EMBL; X62636; CAA44502.1; -
 DR EMBL; X59691; CAA42212.1; -
 DR HSSP; P09651; IUPI; -
 DR FlyBase; Fgn0004237; Hrb7F.
 DR InterPro; IPR000504; RNP_rec_mot.
 DR Pfam; PF0076; RRM; 2.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS0103; RRM; 2.
 DR PROSITE; PS00030; RRM_RNP_1; 2.
 KM RNA-binding; Nuclear protein; Ribonucleoprotein; Repeat;
 KW Alternative splicing.
 FT DOMAIN 24 101 RNA-BINDING (RRM) 1.
 FT DOMAIN 115 192 RNA-BINDING (RRM) 2.
 FT VARSPLIC 315 374 MISSING (IN ISOFORM HRP6.1).
 FT CONFLICT 271 271 S->T (IN REF. 3).
 SQ SEQUENCE 386 AA; 39557 MW; 2036C04D1E3AFD CRC64;
 Query Match 8.8%; Score 191; DB 1; Length 386;
 Best Local Similarity 36.2%; Pred. NO. 0.00013;
 Matches 55; Conservative 8; Mismatches 67; Indels 22; Gaps 6;
 QY 6 GGGRCPEQ-----EGGESAGSGGSAIEGGGGSALAPSPYSGVRRRGAR-GGGR- 59
 DB 202 GGGRCPRAGRGGRGGGDRGGGGGGGNGN--RONGGGMWGAGGCGGGGNGGNGGCGG 259
 QY 60 -GRGNKQAGRGGVCGRGGRGRGRGRGRGRGRPPGSGSLGGDGGCGGCGGSGG 118
 DB 260 GSGGMMNQQGSGGSGGPMNNGGNGGNGGNGGSGGGGNGSGWNGNGGGGGGGGFGN 319
 QY 119 -----CGAPRR-----EPYPPFSSGAGRG 137

DB 320 EYQSGYGGGPPRNSMGNRRAPYSGGCGGG 351
 RESULT 15
 CAZ_DROME STANDARD; PRT; 404 AA.
 AC Q27294; Q24445; Q9VX14;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE RNA-binding protein cabeza (Sarcoma-associated RNA-binding fly
 DE homolog) (P19)
 GN CAZ OR SARFH OR CG3606.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1 AND 2).
 RC STRAIN=Canton-S;
 RX MEDLINE=95349623; PubMed=7623847;
 RA Immanuel D., Zinsner H., Ron D.;
 RT "Association of SARFH (sarcoma-associated RNA-binding fly homolog)
 RL with regions of chromatin transcribed by RNA polymerase II.";
 RL Mol. Cell. Biol. 15:4562-4571(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=Canton-S;
 RX MEDLINE=95223793; PubMed=7708500;
 RA Stolow D.T., Haynes S.R.;
 RT "Cabeza, a Drosophila gene encoding a novel RNA binding protein,
 RL shares homology with EWS and TLS, two genes involved in human sarcoma
 RL formation.";
 RL Nucleic Acids Res. 23:835-843(1995).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikher S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aguayo A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Hochs S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferreira C., Fertler S., Fleischmann W.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heilmann J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mitsuhashi N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiterer K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtikas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 [4]
 RN SEQUENCE OF 39-404 FROM N.A.
 RP STRAIN-Oregon R;
 RA Haynes S.R.;
 RL Submitted (Apr-1988) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 212-261 FROM N.A.
 RC STRAIN-Oregon R;
 RX MEDLINE=87175568; PubMed=3031652;
 RA Haynes S.R., Redbert M.L., Mozer B.A., Forquignon F., David I.B.,
 RT "Pen repeat sequences are GGN clusters and encode a glycine-rich
 domain in a *Drosophila* cDNA homologous to the rat helix destabilizing
 protein,"
 RT Proc. Natl. Acad. Sci. U.S.A. 84:1819-1823(1987).
 CC -1- FUNCTION: MAY PARTICIPATE IN A FUNCTION COMMON TO THE EXPRESSION
 CC OF MOST GENES TRANSCRIBED BY RNA POLYMERASE II.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);
 CC MAY BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS. ENRICHED IN THE BRAIN AND CENTRAL
 CC NERVOUS SYSTEM DURING EMBRYOGENESIS. ENRICHED IN THE ADULT HEAD.
 CC EMBRYOS CONTAIN BOTH TYPE 1 AND TYPE 2 ISOFORMS, WHEREAS LATER IN
 CC DEVELOPMENT (HEADS AND TORSOS) ONLY THE TYPE 2 ISOFORM IS
 CC DETECTED.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE DEVELOPING EMBRYO FROM THE
 CC EARLIEST STAGES OF CELLULARIZATION AND IS SUBSEQUENTLY FOUND IN
 CC MANY CELL TYPES.
 CC -1- MISCELLANEOUS: 'CABEZA' MEANS 'HEAD' IN SPANISH.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNM).
 CC -1- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
 CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U13178; AA86955.1; -
 DR EMBL: L37083; AAC41563.1; -
 DR EMBL: AE003501; AAF48578.1; -
 DR EMBL: M15765; AAT70425.1; -
 DR FLYBase: FBgn0011571; caz.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR InterPro: IPR001876; Znf_RanGDP.
 DR Pfam: PF00076; rtm; 1.
 DR Pfam: PF00641; zf-RanBP; 1.
 DR SMART: SM00360; RRM; 1.
 DR SMART: SM00547; Znf_RBZ; 1.
 DR PROSITE: PS0102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 DR PROSITE: PS01358; ZF_RANBP2_1; 1.
 DR PROSITE: PS01099; ZF_RANBP2_2; 1.
 DR Nuclear protein; zinc-finger; Metal-binding; RNA-binding;
 KM Alternative splicing
 FT DOMAIN 42 111 GLY-RICH.
 FT DOMAIN 119 205 RNA-BINDING (RRM).
 FT DOMAIN 212 275 GLY-RICH.
 FT ZN_FING 280 309 RANBP2-TYPE.
 FT DOMAIN 312 391 GLY-RICH.
 FT VARSPLIC 4 47 MISSING (IN ISOFORM 1).
 FT CONFLICT 39 41 PNY -> LFI (IN REF. 4).
 FT CONFLICT 92 92 P -> H (IN REF. 3).
 FT CONFLICT 108 108 G -> GG (IN REF. 3).
 FT CONFLICT 253 258 MISSING (IN REF. 3).
 FT CONFLICT 283 283 D -> E (IN REF. 4 AND 5).

FT CONFLICT 389 398 DGGPMNDG3 -> MYDDEKRWMS (IN REF. 4).
 SQ SEQUENCE 404 AA; 39141 MM; 7062A0446BA5984 CRC64;
 Query Match 8.8%; Score 190; DB 1; Length 404;
 Best Local Similarity 30.3%; Pred. No. 0.00016;
 Matches 59; Conservative 4; Mismatches 68; Indels 64; Gaps 5;

QY 6 GGGRCCEPDEGESAGSGAGSALIEQGGQSALAPSPVSGVRRGARGGGRGR-- 63
 |||||
 Db 219 GGGRGCGFGRRGG 278
 QY 64 -----WK----- -QAGRGGVCGRGRGRGRGRGRGR 89
 Db 279 PPDGDMKCNSCNNTNFAMRNECNCRKTPKXDDSSSGGGGGGGGGGGGGGGGG 338
 QY 90 RG-----RGRGRPPSGSGGLGDDGGC-----GGGSGGGGAPRRPPVPSGSA 134
 |||||
 Db 339 SGGGGYHNDRGNGSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 389
 QY 135 GPGPRGRATESGKR 149
 |||||
 Db 390 ---GGPMRNDGGM 400

Search completed: March 12, 2003, 08:44:08
 Job time : 28.489 secs

QY	Db	QY	Db
507	-----PTSKSLIVNTIVDS--DMHTGYC-----ERIQ-----	531	
246	ROTASIFKOPVYVHNHNSNKVKYSDPQRNEOPR-QLFPERKLGLSASDVTEQIIKTEME	304	
552	-----VININSPFTWTVTITSIRCNIGTILMSARSRVGNALTLASGMOWN--KTLD	579	
305	LPKGLGVGPGSN 317		
580	-----PKGTTERFGCSN 591		

RESULT 2

glycine-rich protein 1.0 precursor - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C>Date: 30-Sep-1999 #sequence_revision 19-May-1994 #text_change 16-Jul-1999
C/Accession: S01821
R/Keller, B.; Sauer, N.; Lamb, C.J.
EMBO J. 7, 3625-3633, 1988
A>Title: Glycine-rich cell wall proteins in bean: gene structure and association of the
A/Reference number: S01820; MUID:8909j1109; PMID:3208742
A/Accession: S01821
A/Molecule type: DNA
A/Residues: 1-252 <KEI>
A/Cross-references: EMBL:X13595; NID:g21000; PIDN:CAA31931.1; PID:g21001
C/Superfamily: Phaseolus glycyne-rich protein 1.0
C/Keywords: cell wall; structural protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-252/Product: glycine-rich protein 1.0 #status predicted <MAT>

Query Match	10.28;	Score 221.5;	DB 1;	Length 252;
-------------	--------	--------------	-------	-------------

Matches 53; Conservative 7; Mismatches 64; Indels 5; Gaps 2;

QY	17	GESAGGSGAGAGDGAIDEDGGCGSLALSPSVSGVREARCGRRGCRWMKADRGCGVGR	76
		: :	
Dd	87	GHGCGCGNGCGGGGADGGDGGTGGCAGRGCGETGCGGANGSGTGYGGGGSSGGGGGAGAGA	146
QY	77	GRGRGRGRGRGRGRG----RCRCRPSCGSLGDGGCGCGGSGCGG--CAPRRPVPFP	131
		: :	
Dd	147	GSGTGGGEGGSAGCGCYGANCGCGCGGCGGSGSAGAHGGAGAAAGGEGCAHQCAAGGTGG	206
QY	132	GSACGPGRP	140
		: :	
Dd	207	GAAGGGG	215

RESULT 3

glycine-rich α1 wall structural protein homolog F33EJ3.120 α-Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 14-May-1999
C:Accession: J04552
R:Bevan, M.; Hillbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Jesse, T.
Submitted to the Protein Sequence Database, March 1998
#:Reference number: 215378

```

A:Molecule type: DNA
A:Residues: 1-221 <ENV>
A:Cross-references: EMBL:AL022141
A:Experimental source: Cultivar Columbia; BAC clone F23E13
C:Genetics:
A:Map position: 4
A:Note: F23E13.120

```

Query Match	Score	DB 2;	Length
10.28;	220;	221;	

Matches 62; Conservative 5; Mismatches 50; Indels 36; Gaps 5;

```

0Y 15 EEGGS-AAAGSGAGGDSAIEGGGGGSLAPVSGVREGARGCGRG-----RW 64
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 82 EKCESRAGGGGGGGGGGGGGGGGGS-----GGGGGEGNGCKDNSHKRN 127

```

Oy 65 KOLRGAGVCGGRGRRGRRGRRGRRGRRGPSSGLGDGGCCGGSGGGGAPRR 124
 :
Db 128 KSSGGGGG--GGGGGGGGGNGSGRGRGGGGGGGGGGGGGGGGGGGGGGGGDQKG 186
 :
Oy 125 EPVPFSPGSAGRPGR-----PRATES 146
 :
Db 187 GGMFGFGMGDDPGGNFRGCMVRCCTTKKSTES 219

RESULT 4

myosin heavy chain IB - Acanthamoeba castellanii
 M:Contains: myosin ATPase (EC 3.6.4.1)
 C:Species: Acanthamoeba castellanii
 C:Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 19-Apr-2002
 C:Accession: JQ0095; B34448
 R:Jung, G.; Schmidt, C.J.; Hammer III, J.A.
 Gene 82, 269-280, 1989
 A:Title: Myosin I heavy-chain genes of Acanthamoeba castellanii: cloning of a second
 A:Reference number: JQ0095; MUID:90060816; PMID:2511079
 A:Accession: JQ0095
 A:Molecule type: DNA
 A:Residues: 1-1147 <JUN>
 A:Cross-references: GB:M30780
 A:Note: this organism expresses at least three isoforms of myosin I heavy-chain, encoded by gene MIB, whereas the protein previously identified as MIB is the product of gene R.Brzeska, H.; Lynch, T. J.; Mattlin, B.; Korn, E.D.
 J. Biol. Chem. 264, 19340-19348, 1989
 A:Title: The localization and sequence of the phosphorylation sites of Acanthamoeba m

A;Molecule type: protein

C:Comment: In this protein, the coiled-coil rod-like region found in many myosin heavy

protein is globular and does not self-assemble into filaments.

A;Gene: MIB

A:Introns: 1/159/93/3, 102/2, 135/3; 183/3; 212/1; 291/3; 379/3; 492/3; 611/2; 648/3
 C:Superfamily: protozoan myosin heavy chain 1B; myosin motor domain homology; SH3 hom
 C:Keywords: actin binding; ATP; hydrolase; nucleotide binding; P-loop; phosphoprotein
 E:12-664/Domain: myosin motor domain homology <MOT>
 F:103-110/Region: nucleotide-binding motif A (P-loop)
 F:352-513/Region: actin binding #status predicted
 F:698-1147/Domain: carboxyl-terminal <CTD>

F;910-1094/Region: alanine/glycine/proline-rich
F;1097-1144/Domain: SH3 homology <SH3>

F;109/Binding site: ATP (Lys)	#status predicted
F;315/Binding site: phosphate (Ser)	#status predicted

Query Match	10.18;	Score 219.5;	DB 1;	Length 1147;
-------------	--------	--------------	-------	--------------

Matches	81;	Conservative	12;	Mismatches	85;	Indels	47;	Gaps	12;
---------	-----	--------------	-----	------------	-----	--------	-----	------	-----

QY	2	RAHFGGRCRCPEDQEEESNAGSGSAGAGDSALTEGGGGSALAPSVPSSVGRREGARG----	GG	58
Db	926	RGCGGGGGR--GAAGGCGGCGGGGGGGGYSOPVAQAOPVAQVPPAANAAPSAGRCGPGMG		983
QY	59	RGGRNMQA-GRGG-GVCGGRGRCGRGRCGRGRCGRGRCRPPSSGSLGDDG--	CGCG	113
Db	984	PGAGRGCPGKGRGRCPGMGCGAORG-CPGMGCGCGRCRGGPCGPGAGRGRCGPGPAAGRG		1044
QY	114	GGSGG-GGADRREVPFPSSGAGCGRGRPRATESGCRMDCPALPGMKKEVIKSGLSA		172
Db	1043	PGMGPGGAGRGGR--GAORGPGMGCGGAGRG-----	PG-----	A 1077
QY	173	GKSYVYTFPSGKKFKSKPOLAVLYLGNVYDLSFDPRTGKMMPSK		217
Db	1078	GRCAAPAPAPAD--AKPPVKA-----LVDTYAQGTDELTFK		1112

RESULT 5

T20410

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:06 11 November 2014

1

1

OY 76 RGRGRGRGRGRGRGRGRPPSGSGIGDGGGGGGGGGSGGAPRRREPVPFPPSGSAG 135

Db 84 GGNGSGSGSGYGYGYGQNGAGAGQGGGGGGGGGGGGGGSGSGYGYGYGKGGGG 143

OY 136 PGPRG 140

Db 144 GGGGG 148

Search completed: March 12, 2003, 09:13:50
Job time : 39.9242 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2003, 05:39:48 ; Search time 18.9316 Seconds

(without alignments)
915.501 Million cell updates/sec

Title: US-09-554-414b-2

Perfect score: 2167
Sequence: 1 MRAHPGGGRCPEDEEGESGA.....LSRADTEEMDIEMDSGDEA 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PTC_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PTCUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2167	100.0	411	10	US-09-749-728b-1
2	209.5	9.7	201	10	US-09-848-990-22
3	209.5	9.7	201	10	US-09-760-364-14
4	203.5	9.4	357	10	US-09-864-761-35807
5	200.5	9.3	200	9	US-10-160-354-4
6	200.5	9.3	200	9	US-09-990-940-21
7	200.5	9.3	200	9	US-10-026-021-8
8	200.5	9.3	200	9	US-10-161-165-3
9	200.5	9.3	200	9	US-10-160-663-3
10	200.5	9.3	200	10	US-09-798-584-18
11	200.5	9.3	200	10	US-09-967-624-19
12	200.5	9.3	200	10	US-09-998-667-18
13	189.5	8.7	651	10	US-09-861-597-1
14	187.5	8.7	1079	9	US-09-820-843A-20
15	183.5	8.5	283	10	US-09-864-761-36720
16	182.5	8.4	595	9	US-09-854-133-187
17	182.5	8.4	595	10	US-09-738-973-187
18	182.5	8.4	606	10	US-09-861-597-4
19	180.5	8.3	484	9	US-09-820-843A-19

20	178	8.2	645	10	US-09-919-172-41	Sequence 41, Appl
21	176.5	8.1	606	10	US-09-861-597-6	Sequence 6, Appl1
22	175.5	8.1	606	10	US-09-861-597-8	Sequence 8, Appl1
23	174.5	8.1	334	10	US-09-925-301-1363	Sequence 1363, Ap
24	170.5	7.9	256	9	US-09-820-843A-18	Sequence 18, Appl
25	169	7.8	1004	9	US-09-738-676-5676	Sequence 5676, Ap
26	162	7.5	218	10	US-09-925-300-1671	Sequence 1671, Ap
27	159.5	7.4	618	10	US-09-925-300-1381	Sequence 1381, Ap
28	156	7.2	141	10	US-09-864-761-36181	Sequence 36181, A
29	155	7.2	2211	9	US-10-096-961-1	Sequence 1, Appl1
30	151	7.0	714	10	US-09-978-242-3	Sequence 3, Appl1
31	150.5	6.9	101	10	US-09-861-597-3	Sequence 3, Appl1
32	150	6.9	166	10	US-09-837-869A-21	Sequence 21, Appl
33	150	6.9	166	10	US-09-841-321A-21	Sequence 21, Appl
34	147	6.8	173	10	US-09-864-761-36371	Sequence 36371, A
35	146	6.7	191	10	US-09-864-761-36385	Sequence 36385, A
36	145.5	6.7	378	10	US-09-849-967A-2	Sequence 2, Appl1
37	144.5	6.7	440	9	US-10-066-500-106	Sequence 106, App
38	144.5	6.7	440	9	US-10-063-547-52	Sequence 52, App
39	144.5	6.7	440	9	US-10-174-580-202	Sequence 202, App
40	144.5	6.7	440	9	US-10-176-758-202	Sequence 202, App
41	144.5	6.7	440	9	US-10-063-616-52	Sequence 52, Appl
42	144.5	6.7	440	9	US-10-175-737-202	Sequence 202, Appl
43	144.5	6.7	440	9	US-10-063-502-52	Sequence 52, Appl
44	144.5	6.7	440	9	US-10-173-706-202	Sequence 202, App
45	144.5	6.7	440	9	US-10-175-738-202	Sequence 202, App

ALIGNMENTS

RESULT 1
US-09-749-728b-1
Sequence 1, Application US/09749728b
Patent No. US20020142457A1

GENERAL INFORMATION:

APPLICANT: Umezawa, Akihito
APPLICANT: Hata, Jun-ichi
APPLICANT: Fukuda, Keiichi
APPLICANT: Ogawa, Satoshi
APPLICANT: Sakurada, Kazuhiro
APPLICANT: Gojio, Satoshi
APPLICANT: Yamada, Yoji
TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDI
FILE REFERENCE: 00766.000043
CURRENT APPLICATION NUMBER: US/09/749, 728B
CURRENT FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: H11-372826
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: PCT-JP00-01148
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT-JP00-07741
PRIOR FILING DATE: 2000-11-02
NUMBER OF SEQ ID NOS: 80
SOFTWARE: PatentIn Ver.2.0
SEQ ID NO 1
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-09-749-728b-1

Query Match 100.0%; Score 2167; DB 10; Length 411;

Best local similarity 100.0%; Pred. No. 3.2e-144; Mismatches 0; Indels 0; Gaps 0;

Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAHPGGGRCPEDEEGESAGSGAGDSAIIEGGGGSALAPSPVSGVRREGARGRG 60

DB 1 MRAHPGGGRCPEDEEGESAGSGAGDSAIIEGGGGSALAPSPVSGVRREGARGRG 60

QY 61 RGRWKQAGRGCGVCGRGGRGRGRGRGRGRGRGRGRGRGRGRGRGRGRGRGRGR 120

DB 61 RGRWKQAGRGCGVCGRGGRGRGRGRGRGRGRGRGRGRGRGRGRGRGRGRGR 120

```

QY 121 APRREPVFPGSAGPGRGRATESGKRMDCPALPGMKKEEYIRKSGLSAGKSDVYF 180
      |||
Db 121 APRREPVFPGSAGPGRGRATESGKRMDCPALPGMKKEEYIRKSGLSAGKSDVYF 180
QY 181 SPSSKFRKSPOLARYLGNIVDLSFDFRFGKMPKSLQKNKORLNDPLNOKNGKPDNL 240
      |||
Db 181 SPSSKFRKSPOLARYLGNIVDLSFDFRFGKMPKSLQKNKORLNDPLNOKNGKPDNL 240
QY 241 TTLIRQASTFKOPVKVTHNHPNSNKYKSDPORNMEDPROLFWFKRLQGLSASVTEQII 300
      |||
Db 241 TTLIRQASTFKOPVKVTHNHPNSNKYKSDPORNMEDPROLFWFKRLQGLSASVTEQII 300
QY 301 KTMELPKIOLGVGSGNDELTLISAVASALHTSSAPITGOVSAAEKNPAAVWLNTSOPLCR 360
      |||
Db 301 KTMELPKIOLGVGSGNDELTLISAVASALHTSSAPITGOVSAAEKNPAAVWLNTSOPLCR 360
QY 361 AFIVTDEDIRKQERVOOVRRKLEALMADILSRADTEEMDIEMSDGEA 411
      |||
Db 361 AFIVTDEDIRKQERVOOVRRKLEALMADILSRADTEEMDIEMSDGEA 411

```

RESULT 2

```

US-09-848-990-22
; Sequence 22, Application US/09848990
; Patent No. US20020048572A1
; GENERAL INFORMATION:
; APPLICANT: Shan, Bei
; APPLICANT: Schultz, Joshua
; APPLICANT: Tu, Hua
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Treatment of Hypertiglyceridemia and Other Conditions
; FILE REFERENCE: 018781-004910US
; CURRENT APPLICATION NUMBER: US/09/848,990
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 60/201,601
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 22
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: flexible linker
; NAME/KEY: MOD_RES
; LOCATION: (1)..(97)
; OTHER INFORMATION: Gly at positions 1-97 may be present or absent
; NAME/KEY: MOD_RES
; LOCATION: (105)..(201)
; OTHER INFORMATION: Gly at positions 105-201 may be present or absent
US-09-848-990-22

```

```

Query Match          9.7%; Score 209.5; DB 10; Length 201;
Best Local Similarity 40.7%; Pred. No. 9.2e-08;
Matches 55; Conservative 0; Mismatches 63; Indels 17; Gaps 2;

QY 6 GGGRCCEPDEEGESAGSGAGSALIEGGGGSALAPSPVSGYRRRGARCGRGGRGK 65
      |||
Db 17 GGG-----GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 68
QY 66 QAGRGCGVCGRGRGRGRGRGRGRGRGRGRPSGSGSLGDDGGCGGSGGAGPARR 125
      |||
Db 69 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 124
QY 126 PVFPSSAGPGRG 140
      |||
Db 125 -----GGGGGGGGGG 134

```

```

RESULT 3
US-09-760-364-14
; Sequence 14, Application US/09760364

```

```

; Patent No. US20020152479A1
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Juegen Michael
; APPLICANT: Shiau, Andrew Kwan-Nan
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: CAR Modulators: Screening and Treatment of
; FILE REFERENCE: 018781-004110US
; CURRENT APPLICATION NUMBER: US/09/760,364
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/176,398
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 14
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: flexible linker
; NAME/KEY: MOD_RES
; LOCATION: (1)..(97)
; OTHER INFORMATION: Gly residues from positions 1-97 may be present or
; NAME/KEY: MOD_RES
; LOCATION: (105)..(201)
; OTHER INFORMATION: Gly residues from, positions 105-201 may be
; OTHER INFORMATION: present or absent
US-09-760-364-14

```

```

Query Match          9.7%; Score 209.5; DB 10; Length 201;
Best Local Similarity 40.7%; Pred. No. 9.2e-08;
Matches 55; Conservative 0; Mismatches 63; Indels 17; Gaps 2;

```

```

QY 6 GGGRCCEPDEEGESAGSGAGSALIEGGGGSALAPSPVSGYRRRGARCGRGGRGK 65
      |||
Db 17 GGG-----GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 68
QY 66 QAGRGCGVCGRGRGRGRGRGRGRGRGRGRPSGSGSLGDDGGCGGSGGAGPARR 125
      |||
Db 69 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 124
QY 126 PVFPSSAGPGRG 140
      |||
Db 125 -----GGGGGGGGGG 134

```

```

RESULT 4
US-09-864-761-35807
; Sequence 35807, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```

```

CURRENT APPLICATION NUMBER: US/10/160,354
CURRENT FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: US 60/296,819
PRIOR FILING DATE: 2001-06-07
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 200
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: flexible linker
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (6)-(200)
OTHER INFORMATION: Gly at positions 6-200 may be present or absent
US-10-160-354-4

Query Match          9.3%; Score 200.5; DB 9; Length 200;
Best Local Similarity 40.0%; Pred. No. 3.9e-07;
Matches 54; Conservative 0; Mismatches 64; Indels 17; Gaps 2

OY      6 GGGCGCCPEDEGSAMGSGAGSDSAIEQGGGGSALAPVSGVAREGARCGRGRMK 65
        |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      1 GGG-----GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 52
        |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY      66 QAGRGCGVCGRGRNGRNGRNGRNGRGRGRPPSGSSGLGDDGGCGGGGGCGAPRE 125
        |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      53 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 108
        |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY      126 PVPPSGSAGCPRG 140
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      109 -----GGGGGGGGGG 118
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 5
US-09-990-940-21
Sequence 21, Application US/09990940
Publication No. US20030027252A1
GENERAL INFORMATION:
APPLICANT: Tian, Hui
APPLICANT: Zhao, Jiayang
APPLICANT: Chen, Jin-Long
APPLICANT: Cutler, Gene
APPLICANT: An, Songzhu
APPLICANT: Dai, Kang
APPLICANT: Gupta, Jamila S.
APPLICANT: Tularik Inc.
TITLE OF INVENTION: NO. US20030027252A1e1 Receptors
FILE REFERENCE: 018/781-007410US
CURRENT APPLICATION NUMBER: US/09/990,940
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/252,841
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: US 60/257,636
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 60/261,377
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: US 60/279,554
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: US 60/280,696
PRIOR FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 200
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: flexible linker
NAME/KEY: MOD_RES
LOCATION: (6)-(200)
OTHER INFORMATION: Gly at positions 6-200 may be present or absent

```

US-09-990-940-21

Query Match 9.3%; Score 200.5; DB 9; Length 200;
Best Local Similarity 40.0%; Pred. No. 3.9e-07;
Matches 54; Conservative 0; Mismatches 64; Indels 17; Gaps 2;

QY 6 GGGRCPEDEEBSAAGSGAGDSALIEQGGGSLAPSPVSVYREGARGGRGRWK 65
DB 1 GGG-----GG 52
QY 66 QAGRGCGVCGR 125
DB 53 GGG 108
QY 126 PVPFPGSAGPGPRG 140
DB 109 -----GGGGGGGGGG 118

RESULT 7

US-10-026-021-8
Sequence 8, Application US/10026021
Publication No. US2003002756A1
GENERAL INFORMATION:
APPLICANT: Hitoishi, Yasumichi
APPLICANT: Jenkins, Yonchu
TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for
FILE REFERENCE: 021044-001210US
CURRENT APPLICATION NUMBER: US/10/026,021
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/309,632
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 200
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: flexible linker
NAME/KEY: MOD_RES
LOCATION: (6)..(200)
OTHER INFORMATION: Gly at positions 6-200 may be present or absent
US-10-026-021-8

Query Match 9.3%; Score 200.5; DB 9; Length 200;
Best Local Similarity 40.0%; Pred. No. 3.9e-07;
Matches 54; Conservative 0; Mismatches 64; Indels 17; Gaps 2;

QY 6 GGGRCPEDEEBSAAGSGAGDSALIEQGGGSLAPSPVSVYREGARGGRGRWK 65
DB 1 GGG-----GG 52
QY 66 QAGRGCGVCGR 125
DB 53 GGG 108
QY 126 PVPFPGSAGPGPRG 140
DB 109 -----GGGGGGGGGG 118

RESULT 8

US-10-161-165-3
Sequence 3, Application US/10161165
Publication No. US2003002763A1
GENERAL INFORMATION:
APPLICANT: Bennett, Mark
APPLICANT: Holland, Sacha

APPLICANT: Rossi, Alex
APPLICANT: Rigel Pharmaceuticals, Incorporated
TITLE OF INVENTION: CD43: Modulators of Mast Cell Degranulation
FILE REFERENCE: 021044-001010US
CURRENT APPLICATION NUMBER: US/10/161,165
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/296,801
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 200
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: poly Gly tag
NAME/KEY: MOD_RES
LOCATION: (6)..(200)
OTHER INFORMATION: or absent
US-10-161-165-3

Query Match 9.3%; Score 200.5; DB 9; Length 200;
Best Local Similarity 40.0%; Pred. No. 3.9e-07;
Matches 54; Conservative 0; Mismatches 64; Indels 17; Gaps 2;

QY 6 GGGRCPEDEEBSAAGSGAGDSALIEQGGGSLAPSPVSVYREGARGGRGRWK 65
DB 1 GGG-----GG 52
QY 66 QAGRGCGVCGR 125
DB 53 GGG 108
QY 126 PVPFPGSAGPGPRG 140
DB 109 -----GGGGGGGGGG 118

RESULT 9

US-10-160-663-3
Sequence 3, Application US/10160663
Publication No. US20030040001A1
GENERAL INFORMATION:
APPLICANT: Demo, Susan
APPLICANT: Hitoishi, Yasumichi
APPLICANT: Pearsall, Denise
TITLE OF INVENTION: LEMT1: Modulators of Cellular Proliferation
FILE REFERENCE: 021044-000920US
CURRENT APPLICATION NUMBER: US/10/160,663
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/296,817
PRIOR FILING DATE: 2001-06-07
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 200
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: poly Gly tag
NAME/KEY: MOD_RES
LOCATION: (6)..(200)
OTHER INFORMATION: Gly residues from position 6 to 200 may be present
US-10-160-663-3

Query Match 9.3%: Score 200.5; DB 9; Length 200;
Best Local Similarity 40.0%; Pred. No. 3.9e-07;
Matches 54; Conservative 0; Mismatches 64; Indels 17; Gaps 2;

QY 6 GGGRCPEDEGESAGSGAGSDSAIEOGGGSALAPSPVSGVREGARGGRGRWK 65
DB 1 GGG-----GG 52
QY 66 QAGRGCGVCGR 125
DB 53 GGG 108
QY 126 PVFPFSGSAGPGRG 140
DB 109 -----GGGGGGGGGG 118

RESULT 10
US-09-798-584-18
; Sequence 18, Application US/09798584
; Patent No. US20020102676A1
; GENERAL INFORMATION:
; APPLICANT: Mu, David
; APPLICANT: Powers, Scott
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: KCNB: A NO. US20020102676A1el Potassium Channel Protein
; FILE REFERENCE: 018781-004010US
; CURRENT APPLICATION NUMBER: US/09/798,584
; CURRENT FILING DATE: 2001-03-03
; PRIOR APPLICATION NUMBER: US 60/186,951
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:poly-Gly linker
; LOCATION: (6)..(200)
; OTHER INFORMATION: Gly at positions 6-200 may be present or absent
US-09-798-584-18

Query Match 9.3%: Score 200.5; DB 10; Length 200;
Best Local Similarity 40.0%; Pred. No. 3.9e-07;
Matches 54; Conservative 0; Mismatches 64; Indels 17; Gaps 2;

QY 6 GGGRCPEDEGESAGSGAGSDSAIEOGGGSALAPSPVSGVREGARGGRGRWK 65
DB 1 GGG-----GG 52
QY 66 QAGRGCGVCGR 125
DB 53 GGG 108
QY 126 PVFPFSGSAGPGRG 140
DB 109 -----GGGGGGGGGG 118

RESULT 11
US-09-967-624-19
; Sequence 19, Application US/09967624
; Patent No. US20020142325A1
; GENERAL INFORMATION:
; APPLICANT: Liao, X. Charlene
; APPLICANT: Chu, Peter
; APPLICANT: Pardo, Jorge
; APPLICANT: Li, Congfen
; APPLICANT: Zhao, Haoran
; APPLICANT: Wu, Jun

APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: PAK2: Modulators of Lymphocyte Activation
; FILE REFERENCE: 021044-000700US
; CURRENT APPLICATION NUMBER: US/09/967,624
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/280,647
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:poly Gly tag
; FEATURE: Flexible Linker
; NAME/KEY: MOD_RES
; LOCATION: (6)..(200)
; OTHER INFORMATION: Gly residues from position 6 to 200 may be present
; OTHER INFORMATION: or absent
US-09-967-624-19

Query Match 9.3%: Score 200.5; DB 10; Length 200;
Best Local Similarity 40.0%; Pred. No. 3.9e-07;
Matches 54; Conservative 0; Mismatches 64; Indels 17; Gaps 2;

QY 6 GGGRCPEDEGESAGSGAGSDSAIEOGGGSALAPSPVSGVREGARGGRGRWK 65
DB 1 GGG-----GG 52
QY 66 QAGRGCGVCGR 125
DB 53 GGG 108
QY 126 PVFPFSGSAGPGRG 140
DB 109 -----GGGGGGGGGG 118

RESULT 12
US-09-998-667-18
; Sequence 18, Application US/09998667
; Patent No. US20020146747A1
; GENERAL INFORMATION:
; APPLICANT: Masuda, Esteban
; APPLICANT: Liao, X. Charlene
; APPLICANT: Zhao, Haoran
; APPLICANT: Chu, Peter
; APPLICANT: Pardo, Jorge
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: TRAC1: Modulators of Lymphocyte Activation
; FILE REFERENCE: 021044-000600US
; CURRENT APPLICATION NUMBER: US/09/998,667
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 60/282,432
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible linker
; NAME/KEY: MOD_RES
; LOCATION: (6)..(200)
; OTHER INFORMATION: Gly at positions 6-200 may be present or absent
US-09-998-667-18

Query Match 9.3%: Score 200.5; DB 10; Length 200;
Best Local Similarity 40.0%; Pred. No. 3.9e-07;

```

RESULT 15
US-09-864-761-36720
1 Sequence 36720, Application US/09864761
2 Patent No. US20020048763A1
3
4 GENERAL INFORMATION:
5
6 APPLICANT: Penn, Sharon G.
7 APPLICANT: Rank, David R.
8 APPLICANT: Hanzel, David K.
9 APPLICANT: Chen, Wensheng
10 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
11 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
12 FILE REFERENCE: Aecmca-X-1
13 CURRENT APPLICATION NUMBER: US/09/864,761
14 CURRENT FILING DATE: 2001-05-23
15
16 PRIOR APPLICATION NUMBER: US 60/180,312
17 PRIOR FILING DATE: 2000-02-04
18
19 PRIOR APPLICATION NUMBER: US 60/207,456
20 PRIOR FILING DATE: 2000-05-26
21
22 PRIOR APPLICATION NUMBER: US 09/632,366
23 PRIOR FILING DATE: 2000-08-03
24
25 PRIOR APPLICATION NUMBER: GB 24263.6
26 PRIOR FILING DATE: 2000-10-04
27
28 PRIOR APPLICATION NUMBER: US 60/236,359
29 PRIOR FILING DATE: 2000-09-27
30
31 PRIOR APPLICATION NUMBER: PCT/US01/00666
32 PRIOR FILING DATE: 2001-01-30
33
34 PRIOR APPLICATION NUMBER: PCT/US01/00667
35 PRIOR FILING DATE: 2001-01-30
36
37 PRIOR APPLICATION NUMBER: PCT/US01/00664
38 PRIOR FILING DATE: 2001-01-30
39
40 PRIOR APPLICATION NUMBER: PCT/US01/00669
41 PRIOR FILING DATE: 2001-01-30
42
43 PRIOR APPLICATION NUMBER: PCT/US01/00665
44 PRIOR FILING DATE: 2001-01-30
45
46 PRIOR APPLICATION NUMBER: PCT/US01/00668
47 PRIOR FILING DATE: 2001-01-30
48
49 PRIOR APPLICATION NUMBER: PCT/US01/00663
50 PRIOR FILING DATE: 2001-01-30
51
52 PRIOR APPLICATION NUMBER: PCT/US01/00662
53 PRIOR FILING DATE: 2001-01-30
54
55 PRIOR APPLICATION NUMBER: PCT/US01/00661
56 PRIOR FILING DATE: 2001-01-30
57
58 PRIOR APPLICATION NUMBER: PCT/US01/00670
59

```

; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 36720
 ; LENGTH: 283
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC006547.9
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
 ; US-09-864-761-36720

Query Match 8.5%; Score 183.5; DB 10; Length 283;
 Best Local Similarity 37.0%; Pred. No. 9.1e-06;
 Matches 54; Conservative 5; Mismatches 66; Indels 21; Gaps 4;

QY 6 GGGHCCPBOEGESAAAGSGAGGDSAIEOGGGGSALAPSPVSGVRREGARGGGRGRWK 65
 DB 4 GCG-----GGSDGGGGGGGGSDGGGGG-----DGGGGGGSDGGGGDGGGG 48
 QY 66 QAGRGCGVCGRGGRGGRGGRGGRGGRGGRGGRGGRGGRGGRGGRGGRGGRGGR 125
 DB 49 SDGGGGGGGGSDGGGGSDGGGGGGSDGGGGSDGGGGSDGGGGSDGGGGSDGGGGSD 107
 QY 126 PVPPSGSAGPGPRGPRATESGKRM 151
 DB 108 -----GGGGGGGGSDGGGGSDGGGGSD 128

Search completed: March 12, 2003, 09:14:38
 Job time : 25.9316 secs

11
12
13
14

A2858877 19 bp DNA linear GSS 21-FEB-2003
2M0164D14F Mouse 10kb BACIDM library Mus musculus genomic
clone UGCC2M0164D14 F, DNA sequence.
A2858877
A2858877.1 GI:13052498
GSS.
house mouse.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 19)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Petersen,T., Rellly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0164 row: D column: 14
 Seq primer: CGTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 19.
 Location/Qualifiers
 1..19

FEATURES
 source
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U06C2M0164D14"
 /clone_1lb="Mouse 10kb plasmid U06C1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 ORIGIN
 0 a 9 c 10 g 0 t

Query Match
 Best Local Similarity 100.0%; Score 8; DB 17; Length 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||||
 Db 10 CGCGCGCG 17

RESULT 2
 A2858877/c
 LOCUS 19 bp DNA linear GSS 21-FEB-2001
 DEFINITION 2M0164D14F Mouse 10kb plasmid U06C1M library Mus musculus genomic
 ACCESSION A2858877
 VERSION A2858877.1 GI:13052498
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 19)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 CONTACT Robert B. Weiss
 UNIVERSITY of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00

Plate: 0164 row: D column: 14
 Seq primer: CGTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 19.
 Location/Qualifiers
 1..19

FEATURES
 source
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U06C2M0164D14"
 /clone_1lb="Mouse 10kb plasmid U06C1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 ORIGIN
 0 a 9 c 10 g 0 t

Query Match
 Best Local Similarity 100.0%; Score 8; DB 17; Length 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||||
 Db 9 CGCGCGCG 2

RESULT 3
 A2316719
 LOCUS 23 bp DNA linear GSS 29-SEP-2000
 DEFINITION 1M0035A01F Mouse 10kb plasmid U06C1M library Mus musculus genomic
 ACCESSION A2316719
 VERSION A2316719.1 GI:10364814
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 23)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 CONTACT Robert B. Weiss
 UNIVERSITY of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00

Plate: 0035 row: A column: 01
 Seq primer: CGTTGTAACGACGCCACT
 Class: plasmid ends
 High quality sequence stop: 23.
 Location/Qualifiers

FEATURES

1. 23

Source
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U0035A01"
 /clone_11b="Mouse 10kb plasmid U0035A01"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 ORIGIN
 0 a 12 c 11 g 0 t

Query Match
 Best Local Similarity 100.0%; Score 8; DB 17; Length 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 Db 15 CGCGCGCG 22

RESULT 4
 LOCUS A2316719/c 23 bp DNA linear GSS 29-SEP-2000
 DEFINITION 1M0035A01F Mouse 10kb plasmid U0035A01 F, DNA sequence.
 ACCESSION A2316719
 VERSION A2316719.1 GI:10364814
 KEYWORDS GSS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 23)
 Dunn, D., Aoyagi, A., Barber, M., Beacom, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A. and Wright, D., Weis, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00

Plate: 0035 row: A column: 01
 Seq primer: CGTTGTAACGACGCCACT
 Class: plasmid ends
 High quality sequence stop: 23.
 Location/Qualifiers

FEATURES

1. 23

Source
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U0035A01"
 /clone_11b="Mouse 10kb plasmid U0035A01"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 ORIGIN
 0 a 12 c 11 g 0 t

Query Match
 Best Local Similarity 100.0%; Score 8; DB 17; Length 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 Db 10 CGCGCGCG 3

RESULT 5
 LOCUS BH852925 25 bp DNA linear GSS 13-JUN-2002
 DEFINITION SALK_075769.37.15.x Arabidopsis thaliana TDM insertion lines Arabidopsis thaliana genomic clone SALK_075769.37.15.x, DNA sequence.
 ACCESSION BH852925
 VERSION BH852925.1 GI:21423796
 KEYWORDS GSS
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 25)
 Alonso, J.M., Laissie, T.J., Barajas, P., Chen, H., Cheuk, R., Gadgil, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J., and Ecker, J.R.
 A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
 Unpublished (2001)
 Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 538 6379
 Email: ecker@salk.edu
 This is single pass sequence recovered from the left border of

TDNA. This sequence lies within 300 bases of the 5' end of Atg956600.

Class: TDNA tagged.

FEATURES

Source

Location/Qualifiers

1..25

/organism="Arabidopsis thaliana"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_075769.37.15.x"

/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 5 a 6 c 9 g 5 t

ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 25;

Best Local Similarity 100.0%; Pred. No. 3.3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8

DB 13 CGCGCGCG 20

RESULT 6

BH852925/

LOCUS

DEFINITION

Arabidopsis thaliana genomic clone SALK_075769.37.15.x, DNA

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

thale cress.

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 25)

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadriab,

/C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,

/Zimmerman,J. and Ecker,J.R.

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished (2001)

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA. This sequence lies within 300 bases of the 5' end of

Atg956600.

Class: TDNA tagged.

FEATURES

Source

Location/Qualifiers

1..25

/organism="Arabidopsis thaliana"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_075769.37.15.x"

/note="PCR was performed on Arabidopsis thaliana lines"

each of which contains one or more TDNA insertion

elements. The resultant fragment for each line was

directly sequenced to determine the genomic sequence at

the site of insertion. Details of the protocols used can

be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 5 a 6 c 9 g 5 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 8; DB 17; Length 25;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8

DB 20 CGCGCGCG 13

RESULT 7

AZ606831

LOCUS

DEFINITION

clone UGCG1M0428N24 R, DNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 28)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

Plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunne@genetics.utah.edu

Insert Length: 1000 Std Error: 0.00

Plate: 0428 row: N column: 24

Seq primer: CACACGAGAAACAGCTATGACC

Class: Plasmid ends

High quality sequence stop: 28.

FEATURES

Source

Location/Qualifiers

1..28

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UGCG1M0428N24"

/note="Vector: PMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/4nates/>). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PMD42 (G11473211419b1AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 0 a 13 c 15 g 0 t
ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
11111111
Db 13 CGCGCGCG 20

RESULT 8
A2606831 28 bp DNA linear GSS 13-DEC-2000
LOCUS A2606831/c
DEFINITION IM0428N24R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

ACCESSION A2606831
VERSION A2606831.1 GI:11729021

KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 28)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0428 row: N column: 24
Seq primer: CACACAGAAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 28.

FEATURES

source

1. .28
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0428N24"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114[gb]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 0 a 13 c 15 g 0 t
ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
11111111
Db 10 CGCGCGCG 3

RESULT 9
A2627049 28 bp DNA linear GSS 13-DEC-2000
LOCUS A2627049
DEFINITION IM0467E21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

ACCESSION A2627049
VERSION A2627049.1 GI:11749239

KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 28)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE Unpublished (2000)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0467 row: E column: 21
Seq primer: CACACAGAAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 28.

FEATURES

source

1. .28
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0467E21"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (g114732114[gb]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```

BASE COUNT      0 a      14 c      14 g      0 t
ORIGIN
Query Match      100.0%; Score 8; DB 17; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 CGCGCGCG 8
        |||||||
Db      19 CGCGCGCG 26

RESULT 10
LOCUS      AZ627049      28 bp      DNA      linear      GSS 13-DEC-2000
DEFINITION  clone UUCG1M0467E21 R, DNA sequence.
ACCESSION  AZ627049
VERSION    AZ627049.1 GI:11749239
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 28)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
JOURNAL    Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0467 row: E column: 21
            Seq primer: CACACAGCAAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 28.
FEATURES
            source
            1..28
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="UUCG1M0467E21"
            /clone_id="Mouse 10kb plasmid UUCG1M library"
            /sex="Male"
            /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
            /note="Vector: pMD42nv; Purified genomic DNA from M.
            musculus C57BL/6J (male) was obtained from the Jackson
            Laboratory Mouse DNA Resource
            (http://www.jax.org/resources/documents/dnares/). The DNA
            was hydrodynamically sheared by repeated passage through a
            0.005 inch orifice at constant velocity. The sheared DNA
            was blunt end-repaired with T4 DNA polymerase and T4
            polynucleotide kinase. Adaptor oligonucleotides were
            ligated to the blunt ends in high molar excess. The
            adapted DNA was purified and size-selected for a 9.5 to
            10.5 kb range using preparative agarose gel
            electrophoresis. Vector DNA was prepared from a derivative
            of pMD42 (g114732114|gbl|AF129072.1), a copy-number
            inducible derivative of plasmid R1. The vector was ligated
            with adaptors complementary to the insert adaptors and
            purified. The sheared, adapted mouse DNA was annealed to
            adapted vector DNA, and transformed into
            chemically-competent E. coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."

```

```

BASE COUNT      0 a      14 c      14 g      0 t
ORIGIN
Query Match      100.0%; Score 8; DB 17; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 CGCGCGCG 8
        |||||||
Db      10 CGCGCGCG 3

RESULT 11
LOCUS      B1556227      30 bp      mRNA      linear      EST 05-SEP-2001
DEFINITION  603237625F1 NCI_CGAP_Mam3 Mus musculus cDNA IMAGE:5290638 5',
            mRNA sequence.
ACCESSION  B1556227
VERSION    B1556227.1 GI:15443541
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
JOURNAL    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
            cDNA library Preparation: Life Technologies, Inc.
            cDNA library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/ULNL at:
            http://image.llnl.gov
            Plate: LLNL1735 row: a column: 07
            High quality sequence stop: 30.
FEATURES
            source
            1..30
            /organism="Mus musculus"
            /strain="129,C57BL/6J,FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:5290638"
            /clone_id="NCI_CGAP_Mam3"
            /tissue_type="tumor, gross tissue"
            /dev_stage="10 months"
            /lab_host="DH10B"
            /note="Organ: mammary. Vector: PCWV-SPORT6; Site:1; SalI;
            Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
            Library constructed by Life Technologies. Investigators
            providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
            Reference for transgenic model: Xu et al., Nature Genetics
            22, 37-43 (1999)."
BASE COUNT      5 a      8 c      14 g      3 t
ORIGIN
Query Match      100.0%; Score 8; DB 13; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 CGCGCGCG 8
        |||||||
Db      22 CGCGCGCG 29

RESULT 12
LOCUS      B1556227      30 bp      mRNA      linear      EST 05-SEP-2001
DEFINITION  603237625F1 NCI_CGAP_Mam3 Mus musculus cDNA IMAGE:5290638 5',
            mRNA sequence.
ACCESSION  B1556227

```

```

VERSION      BI556227.1 GI:15443541
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    NIH-MGC http://mgi.nci.nih.gov/
TITLE        1 (bases 1 to 30)
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
             Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-remail.nih.gov
             Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLM11735 row: a column: 07
             High quality sequence stop: 30.
FEATURES
source
1..30
/organism="Mus musculus"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5290638"
/clone_lib="NCI_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: Mammary; Vector: PCMV-SPORT6; Site_1: Salivary gland; Site_2: Not; Cloned unidirectionally. Primer: Oligo dt. library constructed by Life Technologies, Inc. providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT      5 a      8 c      14 g      3 t
ORIGIN
Query Match      100.0%; Score 8; DB 13; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGCGCGCG 8
        |||||||
Db      29 CGCGCGCG 22

RESULT 13
BF302851      38 bp      mRNA      linear      EST 21-NOV-2000
LOCUS        602032746F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4167706 5',
DEFINITION   mRNA sequence.
ACCESSION    BF302851
VERSION      BF302851.1 GI:11249409
KEYWORDS
SOURCE       house mouse.
ORGANISM     Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    NIH-MGC http://mgi.nci.nih.gov/
TITLE        1 (bases 1 to 38)
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
             Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-remail.nih.gov
             Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:

```

```

http://image.llnl.gov
Plate: LLM9458 row: 1 column: 11
High quality sequence stop: 38.
FEATURES
source
1..38
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4167706"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: PCMV-SPORT6; Site_1: Not; Site_2: Salivary gland; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      6 a      10 c      21 g      1 t
ORIGIN
Query Match      100.0%; Score 8; DB 12; Length 38;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGCGCGCG 8
        |||||||
Db      30 CGCGCGCG 37

RESULT 14
BF302851/c      38 bp      mRNA      linear      EST 21-NOV-2000
LOCUS        602032746F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4167706 5',
DEFINITION   mRNA sequence.
ACCESSION    BF302851
VERSION      BF302851.1 GI:11249409
KEYWORDS
SOURCE       house mouse.
ORGANISM     Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    NIH-MGC http://mgi.nci.nih.gov/
TITLE        1 (bases 1 to 38)
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
             Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-remail.nih.gov
             Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLM9458 row: 1 column: 11
             High quality sequence stop: 38.
FEATURES
source
1..38
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4167706"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: PCMV-SPORT6; Site_1: Not; Site_2: Salivary gland; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      6 a      10 c      21 g      1 t
ORIGIN
Query Match      100.0%; Score 8; DB 12; Length 38;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CGCGCGCG 8
 |||||||
 Db 37 CGCGCGCG 30

RESULT 15

BM400755 38 bp mRNA linear EST 17-JAN-2002
 LOCUS BM400755
 DEFINITION 5009-0-78-H02.t.1 Chilcoat/Turkewitz cDNA (large fraction)

ACCESSION Tetrahymena thermophila cDNA, mRNA sequence.

VERSION

BM400755
 BM400755.1 GI:18200808

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EST from Tetrahymena thermophila, strain CU428.1, growing cells
 Unpublished (2002)
 Contact: Turkewitz AP
 Molecular Genetics and Cell Biology
 University of Chicago
 920 E. 58th Street, Chicago, IL 60637, USA
 Tel: 773 702 4374
 Fax: 773 702 3172
 Email: apturkew@midway.uchicago.edu
 Seq primer: T3.
 Location/Qualifiers
 1..38
 /organism="Tetrahymena thermophila"
 /strain="CU428.1"
 /db_xref="taxon:5911"
 /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
 /note="Vector: Bluescript SK+; Details on library preparation can be found in Chilcoat and Turkewitz (2001) Proc. Natl. Acad. Sci USA, 98: 8709-8713."

FEATURES

source

BASE COUNT 4 a 12 c 16 g 4 t 2 others
 ORIGIN
 Query Match 100.0%; Score 8; DB 13; Length 38;
 Best Local Similarity 100.0%; Pred. No. 3.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 |||||||
 Db 3 CGCGCGCG 10

RESULT 16

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM400755 38 bp mRNA linear EST 17-JAN-2002
 LOCUS BM400755
 DEFINITION 5009-0-78-H02.t.1 Chilcoat/Turkewitz cDNA (large fraction)
 Tetrahymena thermophila cDNA, mRNA sequence.
 ACCESSION BM400755
 BM400755.1 GI:18200808
 EST.
 Tetrahymena thermophila.
 Tetrahymena thermophila.
 Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
 Hymenostomatida; Tetrahymenina; Tetrahymena.
 1 (bases 1 to 38)
 Turkewitz, A.P., Karrer, K.M., Jahn, C., Orias, E., Kirk, K.E., Frankel, J., and Klobutcher, L.
 EST from Tetrahymena thermophila, strain CU428.1, growing cells
 Unpublished (2002)
 Contact: Turkewitz AP
 Molecular Genetics and Cell Biology
 University of Chicago
 920 E. 58th Street, Chicago, IL 60637, USA
 Tel: 773 702 4374

Fax: 773 702 3172
 Email: apturkew@midway.uchicago.edu
 Seq primer: T3.
 Location/Qualifiers
 1..38
 /organism="Tetrahymena thermophila"
 /strain="CU428.1"
 /db_xref="taxon:5911"
 /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
 /note="Vector: Bluescript SK+; Details on library preparation can be found in Chilcoat and Turkewitz (2001) Proc. Natl. Acad. Sci USA, 98: 8709-8713."

FEATURES

source

BASE COUNT 4 a 12 c 16 g 4 t 2 others
 ORIGIN
 Query Match 100.0%; Score 8; DB 13; Length 38;
 Best Local Similarity 100.0%; Pred. No. 3.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 |||||||
 Db 10 CGCGCGCG 3

RESULT 17 39 bp mRNA linear EST 03-APR-2001
 LOCUS BG538554
 DEFINITION 602567279F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4691923 5',

ACCESSION BG538554
 BG538554.1 GI:13530787
 mRNA sequence.

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 39)
 NIH-MGC http://imgc.ncl.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LICM1511 row: n column: 20
 High quality sequence stop: 39.
 Location/Qualifiers
 1..39
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4691923"
 /clone_lib="NIH_MGC_77"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggcattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor sequence: 5'-ATCTGTGAGGCGCGCGCGCAGC-3' (where B = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

FEATURES

source

BASE COUNT 6 a 13 c 17 g 3 t
 ORIGIN
 Query Match 100.0%; Score 8; DB 12; Length 39;
 Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGCGCGCG 8
|||||||

Db 25 CGCGCGCG 32

RESULT 18
BG538554/c 39 bp mRNA linear EST 03-APR-2001
LOCUS 602567279F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4691923 5',
DEFINITION mRNA sequence.

ACCESSION BG538554
VERSION BG538554.1 GI:13530787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1511 row: n column: 20
High quality sequence stop: 39.
Location/Qualifiers

FEATURES
source
1..39
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4691923"
/clone_1lb="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site: 1;
SfiI (ggcgccgcgcgc); Site: 2; SfiI (ggcgccgcgcgc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCCATTTATGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGCGCGCATG-dT(30)EN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 6 a 13 c 17 g 3 t
ORIGIN

Query Match 100.0%; Score 8; DB 12; Length 39;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGCGCGCG 8
|||||||

Db 32 CGCGCGCG 25

RESULT 19
A2431922 40 bp DNA linear GSS 03-OCT-2000
LOCUS 1M0217D21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0217D21 F, DNA sequence.
ACCESSION A2431922
VERSION A2431922.1 GI:10556031
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
1 (bases 1 to 40)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Bm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0217 row: D column: 21
Seq primer: CGTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 40.
Location/Qualifiers

FEATURES
source
1..40
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0217D21"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (911473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 1 a 20 c 14 g 5 t
ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 40;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGCGCGCG 8
|||||||

Db 24 CGCGCGCG 31

RESULT 20
A2431922 40 bp DNA linear GSS 03-OCT-2000
LOCUS 1M0217D21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0217D21 F, DNA sequence.
ACCESSION A2431922
VERSION A2431922.1 GI:10556031
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus. 1 (bases 1 to 40)
Dunn, D., Royagil, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellily, M., Rose, M., Rose, R., Stokes, R., Tingay, A., von Niederhausern, A. and Wright, D., Meiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

TITLE
JOURNAL
COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0217 row: D column: 21
Seq primer: CGTGTAAACGACGCCACGT
Class: plasmid ends
High quality sequence stop: 40.

FEATURES
source

Location/Qualifiers
1..40
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M021D21"
/clone_lib="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42nv: Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

1 a 20 c 14 g 5 t

Query Match 100.0%; Score 8; DB 17; Length 40;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 31 CGCGCGCG 24

RESULT 21
LOCUS BH796426 40 bp DNA linear GSS 25-APR-2002
DEFINITION 1008094B12.1EL_x1 1008 - RescueMu Grid I Zea mays genomic, DNA
sequence.
ACCESSION BH796426
VERSION BH796426.1 GI:20305035
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 40)
Walbot, V.

Maize genomic sequences found using engineered RescueMu transposon

TITLE
JOURNAL
COMMENT

Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1008094 row: 26
Class: transposon-tagged.

FEATURES
source

Location/Qualifiers
1..40
/organism="Zea mays"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_lib="1008 - RescueMu Grid I"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site www.zmdb.iastate.edu and follow the links for "RescueMu." Grid I was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT
ORIGIN

2 a 15 c 13 g 10 t

Query Match 100.0%; Score 8; DB 17; Length 40;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 28 CGCGCGCG 35

RESULT 22
LOCUS BH796426/c 40 bp DNA linear GSS 25-APR-2002
DEFINITION 1008094B12.1EL_x1 1008 - RescueMu Grid I Zea mays genomic, DNA
sequence.
ACCESSION BH796426
VERSION BH796426.1 GI:20305035
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 40)
Walbot, V.
Maize genomic sequences found using engineered RescueMu transposon
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227

Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1008094 row: 26
Class: transposon-tagged.

FEATURES

source

1. 40
/organism="Zea mays"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_lib="1008 - Rescuemu Grid I"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: Rescuemu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site www.zmdb.iastate.edu and follow the links for 'Rescuemu.' Grid I was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT

2 a 15 c 13 g 10 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 8; DB 17; Length 40;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8

|||||

Db 35 CGCGCGCG 28

RESULT 23

BH639962

LOCUS

1008033A02.2EL_y1 1008 - Rescuemu Grid I Zea mays genomic, DNA

DEFINITION

sequence.

ACCESSION

BH639962

VERSION

BH639962.1

KEYWORDS

GSS.

SOURCE

Zea mays.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

REFERENCE

Walbot, V.

AUTHORS

Unpublished (2001)

TITLE

Contact: Walbot V

JOURNAL

Department of Biological Sciences

COMMENT

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.

Reverse complemented post-ligation sequence from source sequence.

Plate: 1008033 row: 34

Class: transposon-tagged.

Location/Qualifiers

1. 41

/organism="Zea mays"

/cultivar="mixed background W23/A188/B73"

/db_xref="taxon:4577"

/clone_lib="1008 - Rescuemu Grid I"

BASE COUNT

6 a 12 c 21 g 2 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 8; DB 17; Length 41;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8

|||||

Db 15 CGCGCGCG 22

RESULT 24

BH639962/c

LOCUS

1008033A02.2EL_y1 1008 - Rescuemu Grid I Zea mays genomic, DNA

DEFINITION

sequence.

ACCESSION

BH639962

VERSION

BH639962.1

KEYWORDS

GSS.

SOURCE

Zea mays.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

REFERENCE

Walbot, V.

AUTHORS

Unpublished (2001)

TITLE

Contact: Walbot V

JOURNAL

Department of Biological Sciences

COMMENT

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.

Reverse complemented post-ligation sequence from source sequence.

Plate: 1008033 row: 34

Class: transposon-tagged.

Location/Qualifiers

1. 41

/organism="Zea mays"

/cultivar="mixed background W23/A188/B73"

/db_xref="taxon:4577"

/clone_lib="1008 - Rescuemu Grid I"

/tissue_type="leaf"

/dev_stage="adult"

/lab_host="DH10B"

/note="Organ: leaf; Vector: Rescuemu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site www.zmdb.iastate.edu and follow the links for 'Rescuemu.' Grid I was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI

and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 6 a 12 c 21 g 2 t

Query Match 100.0%; Score 8; DB 17; Length 41;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|||||||
DB 22 CGCGCGCG 15

RESULT 25
LOCUS BG915507 43 bp mRNA linear EST 05-JUN-2001
DEFINITION 602815734F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4945113 5', mRNA sequence.
ACCESSION BG915507
VERSION BG915507.1 GI:14295983
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 43)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
Plate: LLM10891 row: d column: 10
High quality sequence stop: 40.
Location/Qualifiers

FEATURES
Source 1..43
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4945113"
/clone_lib="NCI_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

BASE COUNT 4 a 16 c 19 g 4 t

Query Match 100.0%; Score 8; DB 13; Length 43;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|||||||
DB 30 CGCGCGCG 37

RESULT 26
BG915507/o

LOCUS BG915507 43 bp mRNA linear EST 05-JUN-2001
DEFINITION 602815734F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4945113 5', mRNA sequence.
ACCESSION BG915507
VERSION BG915507.1 GI:14295983
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 43)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
Plate: LLM10891 row: d column: 10
High quality sequence stop: 40.
Location/Qualifiers

FEATURES
Source 1..43
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4945113"
/clone_lib="NCI_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

BASE COUNT 4 a 16 c 19 g 4 t

Query Match 100.0%; Score 8; DB 13; Length 43;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|||||||
DB 37 CGCGCGCG 30

RESULT 27
LOCUS BH629051 44 bp DNA linear GSS 30-JAN-2002
DEFINITION 1007076807.2EL_y1 1007 - RescueMu Grid H Zea mays genomic, DNA sequence.
ACCESSION BH629051
VERSION BH629051.1 GI:18442302
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 44)
Walbot,V.
Maize genomic sequences found using engineered RescueMu transposon
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007076 column: 36
Class: transposon-tagged

FEATURES

source

1..44
/organism="Zea mays"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_id="1007 - Rescuemu Grid H"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: Rescuemu (engineered from pluescript backbone); Site: 1: BamHI; Site: 2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'Rescuemu.' Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 15 a 20 c 6 g 3 t

ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 44;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
Db 8 CGCGCGCG 15

RESULT 28
BH629051/c 44 bp DNA linear GSS 30-JAN-2002
LOCUS 1007076B07.2EL_Y1 1007 - Rescuemu Grid H Zea mays genomic, DNA
DEFINITION
sequence.

ACCESSION BH629051.1 GI:18442302
VERSION
KEYWORDS
SOURCE

ORGANISM

Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 44)

REFERENCE
AUTHORS Walbot, V.
TITLE Maize genomic sequences found using engineered Rescuemu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007076 column: 36
Class: transposon-tagged.

FEATURES

source

1..44
/organism="Zea mays"

/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_id="1007 - Rescuemu Grid H"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: Rescuemu (engineered from pluescript backbone); Site: 1: BamHI; Site: 2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'Rescuemu.' Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 15 a 20 c 6 g 3 t

ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 44;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
Db 15 CGCGCGCG 8

RESULT 29
A2514564 45 bp DNA linear GSS 05-OCT-2000
LOCUS IM0361N10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0361N10 F, DNA sequence.
ACCESSION A2514564
VERSION A2514564.1 GI:10695796
KEYWORDS
SOURCE

ORGANISM

house mouse.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 45)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamll, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
'M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)

JOURNAL
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0361 row: N column: 10
Seq primer: CGTTGTAAACGACGCGCACT
Class: plasmid ends
High quality sequence stop: 45.

FEATURES

source

1..45
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_id="UUGC1M0361N10"
/clone_id="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMDA2nv, Purified genomic DNA from M."

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 17 a 22 c 6 g 0 t
ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 45;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 36 CGCGCGCG 43

RESULT 30
LOCUS AZ514564/c 45 bp DNA linear GSS 05-OCT-2000
DEFINITION IM0361N10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0361N10 F, DNA sequence.
ACCESSION AZ514564
VERSION AZ514564.1 GI:10695796
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,K., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0361 row: N column: 10
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 45.
Location/Qualifiers

FEATURES
SOURCE 1. 45
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0361N10"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 17 a 22 c 6 g 0 t
ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 45;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 45 CGCGCGCG 38

RESULT 31
LOCUS AZ400633 47 bp DNA linear GSS 03-OCT-2000
DEFINITION IM0167F06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0167F06 F, DNA sequence.
ACCESSION AZ400633
VERSION AZ400633.1 GI:10515707
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,K., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0167 row: F column: 06
Seq primer: CGTTGTAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 47.
Location/Qualifiers

FEATURES
SOURCE 1. 47
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0167F06"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (g1473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT
ORIGIN
2 a 15 c 23 g 7 t

Query Match 100.0%; Score 8; DB 17; Length 47;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|||||||
Db 15 CGCGCGCG 22

RESULT 32
A2400633 47 bp DNA linear GSS 03-OCT-2000
LOCUS 1M0167F06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0167F06 F, DNA sequence.
ACCESSION A2400633
VERSION A2400633.1 GI:10515707
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 47)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0167 row: F column: 06
Seq primer: CGTTGTAACGACGCCACGT
Class: plasmid ends
High quality sequence stop: 47.
Location/Qualifiers
1..47
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0167F06"
/clone_1id="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD2mv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (g1473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT
ORIGIN
2 a 15 c 23 g 7 t

Query Match 100.0%; Score 8; DB 17; Length 47;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|||||||
Db 24 CGCGCGCG 17

RESULT 33
A2585161 47 bp DNA linear GSS 13-DEC-2000
LOCUS 1M0390G12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0390G12 F, DNA sequence.
ACCESSION A2585161
VERSION A2585161.1 GI:11706772
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 47)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0390 row: G column: 12
Seq primer: CGTTGTAACGACGCCACGT
Class: plasmid ends
High quality sequence stop: 47.
Location/Qualifiers
1..47
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0390G12"
/clone_1id="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD2mv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114[9b]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 14 a 22 c 11 g 0 t
ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 47;

Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||

Db 2 CGCGCGCG 9

RESULT 34

AZ585161/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

AZ585161 47 bp DNA linear GSS 13-DEC-2000
1M0390G12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0390G12 F, DNA sequence.

AZ585161
A2585161.1 GI:11706772

GSS.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 47)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0390 row: G column: 12

Seq primer: CGTTGTAACGACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 47.

Location/Qualifiers

1..47

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0390G12"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114[9b]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 14 a 22 c 11 g 0 t
ORIGIN

Query Match 100.0%; Score 8; DB 17; Length 47;

Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||

Db 9 CGCGCGCG 2

RESULT 35

AZ764402

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

AZ764402 48 bp DNA linear GSS 16-FEB-2001
1M0560018F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0560018 F, DNA sequence.

AZ764402
A2764402.1 GI:12879326

GSS.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 48)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0560 row: 0 column: 18

Seq primer: CGTTGTAACGACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 48.

Location/Qualifiers

1..48

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0560018"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|g9|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
1 a 10 c 24 g 13 t

Query Match 100.0%; Score 8; DB 17; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
Db 2 CGCGCGCG 9

RESULT 36
A2764402/c 48 bp DNA linear GSS 16-FEB-2001
LOCUS 1M0560018F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0560018 F, DNA sequence.
ACCESSION A2764402
VERSION A2764402.1 GI:12879326
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

REFERENCE 1 (bases 1 to 48)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0560 row: 0 column: 18
Seq primer: CCTGTAAACACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 48.
Location/Qualifiers
1..48

FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0560018"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M."

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|g9|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
1 a 10 c 24 g 13 t

Query Match 100.0%; Score 8; DB 17; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
Db 9 CGCGCGCG 2

RESULT 37
AA948394 49 bp mRNA linear EST 23-JUN-1998
LOCUS on52b09.s1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:1560281 3'
DEFINITION similar to SW:CA11.CHICK P02457 PROCOLLAGEN ALPHA 1(I) CHAIN
PRECURSOR.; mRNA sequence.
ACCESSION AA948394
VERSION AA948394.1 GI:3109647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 49)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/URL at:
www.bio.linn.gov/dbp/image/image.html

Trace considered overall poor quality
Insert Length: 1085 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..49

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1560281"
/clone_1lb="NCI_CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: p773D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from

colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 5 a 22 c 18 g 4 t
 Query Match 100.0%; Score 8; DB 9; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 38
 AA948394/c 49 bp mRNA linear EST 23-JUN-1998
 LOCUS 0152009.s1 NCI-CGAP-C08 Homo sapiens cDNA clone IMAGE:1560281 3'
 DEFINITION similar to SW:CA11-CHICK P02457 PROCOLLAGEN ALPHA 1(I) CHAIN
 PRECURSOR. ; mRNA sequence.

ACCESSION AA948394.1 GI:3109647
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 49)
 NCBI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.lnl.gov/dbtrp/image/image.html

Trace considered overall poor quality
 Insert Length: 1085 Std Error: 0.00
 Seq primer: -40ml3 fwd. EP from Amersham
 High quality sequence stop: 1.

FEATURES

SOURCE

1..49
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1560281"
 /clone_lib="NCI-CGAP C08"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"
 /note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 5 a 22 c 18 g 4 t

Query Match 100.0%; Score 8; DB 9; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 Db 38 CGCGCGCG 31

RESULT 39
 A1440059 49 bp mRNA linear EST 09-MAR-1999
 LOCUS t161f10.x1 NCI-CGAP-Lym12 Homo sapiens cDNA clone IMAGE:2134987 3'
 DEFINITION similar to SW:PRPL_HUMAN P10162 SALIVARY PROLINE-RICH PROTEIN PO
 ; contains element TAR1 repetitive element ; mRNA sequence.

ACCESSION A1440059.1 GI:4308506
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 49)
 NCBI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Life Technologies catalog #: 11547-015
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.lnl.gov/dbtrp/image/image.html

Trace considered overall poor quality
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.

FEATURES

SOURCE

1..49
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2134987"
 /clone_lib="NCI-CGAP Lym12"
 /tissue_type="lymphoma, follicular mixed small and large cell"
 /lab_host="DH10B"
 /note="Organ: lymph node; Vector: PCMV-SPORE6; Site: 1; Salt: 2; NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.25 kb. Life Technologies catalog #: 11547-015"

BASE COUNT 13 a 20 c 14 g 2 t

Query Match 100.0%; Score 8; DB 9; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 40
 A1440059/c 49 bp mRNA linear EST 09-MAR-1999
 LOCUS t161f10.x1 NCI-CGAP-Lym12 Homo sapiens cDNA clone IMAGE:2134987 3'
 DEFINITION similar to SW:PRPL_HUMAN P10162 SALIVARY PROLINE-RICH PROTEIN PO
 ; contains element TAR1 repetitive element ; mRNA sequence.

ACCESSION A1440059.1 GI:4308506
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 49)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 JOURNAL Tumor Gene Index
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Life Technologies catalog #: 11547-015
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbcp/image/image.html

FEATURES
 source 1..49
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2134987"
 /clone_1lb="NCI-CGAP_Lym12"
 /tissue_type="lymphoma, follicular mixed small and large
 cell"
 /lab_host="DH10B"
 /note="Organ: lymph node; Vector: pCMV-SPORT6; Site: 1;
 SalI; Site 2: NotI; Cloned unidirectionally. Primer:
 Oligo dt. Average insert size 1.25 kb. Life Technologies
 catalog #: 11547-015"

BASE COUNT 13 a 20 c 14 g 2 t

ORIGIN

Query Match 100.0%; Score 8; DB 9; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCCGC 8
 |||||
 Db 31 CGCGCCGC 24

RESULT 41
 A1565007 49 bp mRNA linear EST 13-MAY-1999
 LOCUS tq53d09.x1 NCI-CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2212529 3'
 DEFINITION similar to TR:Q04154 Q04154 SALIVARY PROLINE-RICH PROTEIN RP15
 PRECURSOR. ; contains PTR5.T3 MSRI repetitive element ;, mRNA
 sequence.
 ACCESSION A1565007
 VERSION A1565007.1 GI:4523464
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 49)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbcp/image/image.html

JOURNAL COMMENT
 Trace considered overall poor quality
 Insert Length: 366 Std Error: 0.00

Seq primer: -40UP from G1bco
 High quality sequence stop: 1
 POLYA-No.

FEATURES
 source 1..49
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2212529"
 /clone_1lb="NCI-CGAP_Ut1"
 /tissue_type="well-differentiated endometrial
 adenocarcinoma, 7 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Life Technologies catalog #:
 11538-014"

BASE COUNT 10 a 21 c 18 g 0 t

ORIGIN

Query Match 100.0%; Score 8; DB 9; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCCGC 8
 |||||
 Db 2 CGCGCCGC 9

RESULT 42
 A1565007 49 bp mRNA linear EST 13-MAY-1999
 LOCUS tq53d09.x1 NCI-CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2212529 3'
 DEFINITION similar to TR:Q04154 Q04154 SALIVARY PROLINE-RICH PROTEIN RP15
 PRECURSOR. ; contains PTR5.T3 MSRI repetitive element ;, mRNA
 sequence.
 ACCESSION A1565007
 VERSION A1565007.1 GI:4523464
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 49)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbcp/image/image.html

JOURNAL COMMENT
 Trace considered overall poor quality
 Insert Length: 366 Std Error: 0.00
 Seq primer: -40UP from G1bco
 High quality sequence stop: 1
 POLYA-No.

FEATURES
 source 1..49
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2212529"
 /clone_1lb="NCI-CGAP_Ut1"
 /tissue_type="well-differentiated endometrial
 adenocarcinoma, 7 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; SalI;

Site-2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

BASE COUNT
ORIGIN

10 a 21 c 18 g 0 t

Query Match 100.0%; Score 8; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
DB 9 CGCGCGCG 2

RESULT 43

AZ456464 49 bp DNA linear GSS 04-OCT-2000
LOCUS 1M0259L12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0259L12 F, DNA sequence.

ACCESSION AZ456464
VERSION AZ456464.1 GI:10614505
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 49)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)

JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0259 row: L column: 12
Seq primer: CGTTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 49.

FEATURES
Location/Qualifiers

1..49

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0259L12"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

15 a 24 c 9 g 1 t

Query Match 100.0%; Score 8; DB 17; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
DB 41 CGCGCGCG 48

RESULT 44

AZ456464 49 bp DNA linear GSS 04-OCT-2000
LOCUS 1M0259L12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0259L12 F, DNA sequence.

ACCESSION AZ456464
VERSION AZ456464.1 GI:10614505
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 49)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)

JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0259 row: L column: 12
Seq primer: CGTTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 49.

FEATURES
Location/Qualifiers

1..49

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0259L12"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to

BASE COUNT 15 a 24 c 9 g 1 t
 ORIGIN
 Query Match 100.0%; Score 8; DB 17; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 DB 46 CGCGCGCG 39

RESULT 45
 A2773338

LOCUS A2773338 49 bp DNA linear GSS 16-FEB-2001
 DEFINITION clone UUGC1M0584L13 R, DNA sequence.

ACCESSION A2773338

VERSION A2773338.1 GI:12897597

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 49)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weis, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0584 row: L column: 13
 Seq primer: CACACGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 49.
 Location/Qualifiers
 1. 49

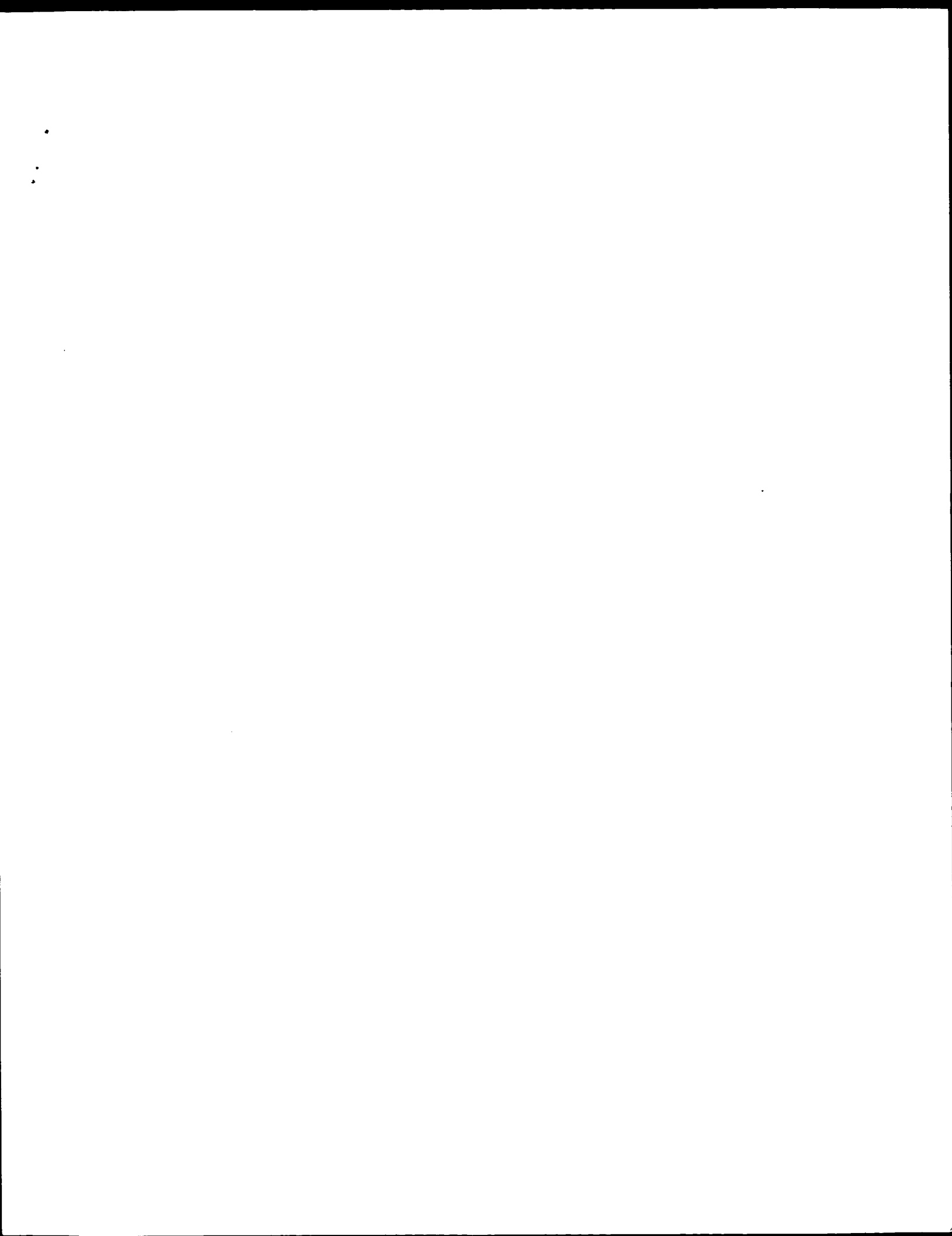
FEATURES
 source

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0584L13"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PMD42 (G114732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptor complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to

BASE COUNT 2 a 8 c 24 g 15 t
 ORIGIN
 Query Match 100.0%; Score 8; DB 17; Length 49;
 Best Local Similarity 100.0%; Pred. No. 2.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 DB 24 CGCGCGCG 31

Search completed: March 14, 2003, 04:32:27
 Job time : 1341 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 03:39:03 ; Search time 60 Seconds
(without alignments)
93.582 Million cell updates/sec

Title: CGCGCGCG
Perfect score: 8
Sequence: 1 cgcgcgcg 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_MA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	100.0	19	9	US-09-888-326-342
2	8	100.0	19	9	US-09-888-326-342
3	8	100.0	20	9	US-09-888-326-192
4	8	100.0	20	9	US-09-888-326-192
5	8	100.0	20	9	US-09-888-326-193
6	8	100.0	20	9	US-09-888-326-193
7	8	100.0	20	9	US-09-888-326-193
8	8	100.0	20	9	US-09-888-326-607
9	8	100.0	21	10	US-09-885-441-54
10	8	100.0	21	10	US-09-885-441-54
11	8	100.0	25	10	US-09-728-721-68
12	8	100.0	25	10	US-09-728-721-68
13	8	100.0	25	10	US-09-841-8798-16
14	8	100.0	25	10	US-09-841-8798-16
15	8	100.0	30	10	US-09-682-597-4
16	8	100.0	30	10	US-09-682-597-4
17	8	100.0	36	10	US-09-828-034-5
18	8	100.0	36	10	US-09-828-034-5
19	8	100.0	39	9	US-09-733-042-31

Result No.	Score	Query Match	Length	ID	Description
20	8	100.0	39	9	US-09-733-042-31
21	8	100.0	45	9	US-10-007-132-45
22	8	100.0	45	9	US-10-007-132-45
23	8	100.0	55	10	US-09-790-417-19
24	8	100.0	55	10	US-09-790-417-19
25	8	100.0	60	10	US-09-899-381-37
26	8	100.0	60	10	US-09-899-381-37
27	8	100.0	61	12	US-10-014-973A-19
28	8	100.0	61	12	US-10-014-973A-19
29	8	100.0	64	9	US-10-057-940-8
30	8	100.0	64	9	US-10-057-940-8
31	8	100.0	68	10	US-09-473-872-54
32	8	100.0	68	10	US-09-473-872-54
33	8	100.0	72	9	US-09-835-9768-122
34	8	100.0	72	9	US-09-835-9768-122
35	8	100.0	74	9	US-09-835-9768-122
36	8	100.0	74	9	US-09-835-9768-119
37	8	100.0	76	9	US-09-835-9768-127
38	8	100.0	76	9	US-09-835-9768-127
39	8	100.0	76	9	US-09-835-9768-129
40	8	100.0	76	9	US-09-835-9768-129
41	8	100.0	78	9	US-09-835-9768-126
42	8	100.0	78	9	US-09-835-9768-126
43	8	100.0	78	9	US-09-835-9768-128
44	8	100.0	78	9	US-09-835-9768-128
45	8	100.0	80	9	US-09-835-9768-124

ALIGNMENTS

RESULT 1

US-09-888-326-342

Sequence 342, Application US/09888326

Publication No. US20030026801A1

GENERAL INFORMATION:

APPLICANT: Weimer, George

APPLICANT: Hartmann, Gunther

TITLE OF INVENTION: Methods for Enhancing Antibody-Induced

FILE REFERENCE: C1039/7052 (AWS)

CURRENT FILING DATE: 2001-06-22

PRIOR APPLICATION NUMBER: US 60/213,346

PRIOR FILING DATE: 2000-06-22

NUMBER OF SEQ ID NOS: 848

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 342

LENGTH: 19

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic oligonucleotide

NAME/KEY: misc:feature

LOCATION: (0)...(0)

OTHER INFORMATION: phosphodiester backbone

US-09-888-326-342

Query Match

Best Local Similarity 100.0%; Score 8; DB 9; Length 19;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 10 CGCGCGCG 17

CGCGCGCG 8

RESULT 2

US-09-888-326-342/c

Sequence 342, Application US/09888326

Publication No. US20030026801A1

GENERAL INFORMATION:

APPLICANT: Weimer, George

```
; APPLICANT: Hartmann, Gunther
; TITLE OF INVENTION: Methods for Enhancing Antibody-Induced
; FILE REFERENCE: C1039/7052 (AMS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 342
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphodiester backbone
US-09-888-326-342
```

```
Query Match          100.0%; Score 8; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CGCGCGCG 8
    |||||||
Db 9 CGCGCGCG 2
```

```
RESULT 3
US-09-888-326-192
; Sequence 192, Application US/09888326
; Publication No. US20030026801A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Hartmann, Gunther
; TITLE OF INVENTION: Methods for Enhancing Antibody-Induced
; FILE REFERENCE: C1039/7052 (AMS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 192
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphorothioate backbone
US-09-888-326-192
```

```
Query Match          100.0%; Score 8; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CGCGCGCG 8
    |||||||
Db 11 CGCGCGCG 18
```

```
RESULT 4
US-09-888-326-192/c
; Sequence 192, Application US/09888326
; Publication No. US20030026801A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Hartmann, Gunther
```

```
; TITLE OF INVENTION: Methods for Enhancing Antibody-Induced
; FILE REFERENCE: C1039/7052 (AMS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 192
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphorothioate backbone
US-09-888-326-192
```

```
Query Match          100.0%; Score 8; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CGCGCGCG 8
    |||||||
Db 10 CGCGCGCG 3
```

```
RESULT 5
US-09-888-326-193
; Sequence 193, Application US/09888326
; Publication No. US20030026801A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Hartmann, Gunther
; TITLE OF INVENTION: Methods for Enhancing Antibody-Induced
; FILE REFERENCE: C1039/7052 (AMS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 193
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphodiester backbone
US-09-888-326-193
```

```
Query Match          100.0%; Score 8; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CGCGCGCG 8
    |||||||
Db 11 CGCGCGCG 18
```

```
RESULT 6
US-09-888-326-193/c
; Sequence 193, Application US/09888326
; Publication No. US20030026801A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Hartmann, Gunther
; TITLE OF INVENTION: Methods for Enhancing Antibody-Induced
```

```
; TITLE OF INVENTION: Cell Lysis and Treating Cancer
; FILE REFERENCE: C1039/7052 (AMS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 193
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphodiester backbone
US-09-888-326-193
```

```
Query Match          100.0%; Score 8; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
OY      1 CGCGCGCG 8
         |||||
Db       10 CGCGCGCG 3
```

```
RESULT 7
US-09-888-326-607
; Sequence 607, Application US/09888326
; Publication No. US20030026801A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Hartmann, Gunther
; TITLE OF INVENTION: Methods for Enhancing Antibody-Induced
; FILE REFERENCE: C1039/7052 (AMS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 607
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphodiester backbone
US-09-888-326-607
```

```
Query Match          100.0%; Score 8; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
OY      1 CGCGCGCG 8
         |||||
Db       13 CGCGCGCG 20
```

```
RESULT 8
US-09-888-326-607/c
; Sequence 607, Application US/09888326
; Publication No. US20030026801A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Hartmann, Gunther
; TITLE OF INVENTION: Methods for Enhancing Antibody-Induced
; FILE REFERENCE: C1039/7052 (AMS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 193
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphodiester backbone
US-09-888-326-607/c
```

```
; FILE REFERENCE: C1039/7052 (AMS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 607
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphodiester backbone
US-09-888-326-607
```

```
Query Match          100.0%; Score 8; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
OY      1 CGCGCGCG 8
         |||||
Db       20 CGCGCGCG 13
```

```
RESULT 9
US-09-885-441-54
; Sequence 54, Application US/09885441
; Patent No. US20020146407A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Eosinophil Serine
; FILE REFERENCE: 04974.00512
; CURRENT APPLICATION NUMBER: US/09/885,441
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/212,844
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/244,171
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: US 60/279,766
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: PCT/
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-885-441-54
```

```
Query Match          100.0%; Score 8; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
OY      1 CGCGCGCG 8
         |||||
Db       6 CGCGCGCG 13
```

```
RESULT 10
US-09-885-441-54/c
; Sequence 54, Application US/09885441
; Patent No. US20020146407A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Eosinophil Serine
; FILE REFERENCE: 04974.00512
; CURRENT APPLICATION NUMBER: US/09/885,441
```

```

; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/212,844
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/244,171
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: US 60/279,766
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: PCT/
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-885-441-54

Query Match          100.0%; Score 8; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
    |||||
Db 13 CGCGCGCG 6

RESULT 11
US-09-728-721-68
; Sequence 68, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-728-721-68

Query Match          100.0%; Score 8; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
    |||||
Db 16 CGCGCGCG 23

RESULT 12
US-09-728-721-68/c
; Sequence 68, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 25
; TYPE: DNA
```

```

; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-728-721-68

Query Match          100.0%; Score 8; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
    |||||
Db 25 CGCGCGCG 18

RESULT 13
US-09-841-879B-16
; Sequence 16, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-841-879B-16

Query Match          100.0%; Score 8; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
    |||||
Db 16 CGCGCGCG 23

RESULT 14
US-09-841-879B-16/c
; Sequence 16, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 25
; TYPE: DNA
```

; ORGANISM: Homo sapiens
US-09-841-8798-16

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 25 CGCGCGCG 18

RESULT 15

US-09-682-597-4
; Sequence 4, Application US/09682597
; Patent No. US20020062503A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Chen, Gullian
; APPLICANT: Hironaka, Catherine
; APPLICANT: Zhou, Hua-Ping
; TITLE OF INVENTION: Glyphosate Tolerant Wheat Plant 33391 and Compositions and Method
; FILE REFERENCE: 38-21(5232)A
; CURRENT APPLICATION NUMBER: US/09/682,597
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-682-597-4

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 3 CGCGCGCG 10

RESULT 16

US-09-682-597-4/c
; Sequence 4, Application US/09682597
; Patent No. US20020062503A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Chen, Gullian
; APPLICANT: Hironaka, Catherine
; APPLICANT: Zhou, Hua-Ping
; TITLE OF INVENTION: Glyphosate Tolerant Wheat Plant 33391 and Compositions and Method
; FILE REFERENCE: 38-21(5232)A
; CURRENT APPLICATION NUMBER: US/09/682,597
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-682-597-4

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 10 CGCGCGCG 3

RESULT 17
US-09-828-034-5
; Sequence 5, Application US/09828034
; Patent No. US20020064771A1
; GENERAL INFORMATION:
; APPLICANT: Zhong, Weidong
; APPLICANT: Hong, Zhi
; APPLICANT: Ferrari, Eric
; TITLE OF INVENTION: HCV REPLICASE COMPLEXES
; FILE REFERENCE: IN01165
; CURRENT APPLICATION NUMBER: US/09/828,034
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: U.S. 60/195,852
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 36
; TYPE: RNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic RNA
US-09-828-034-5

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 27 CGCGCGCG 34

RESULT 18

US-09-828-034-5/c
; Sequence 5, Application US/09828034
; Patent No. US20020064771A1
; GENERAL INFORMATION:
; APPLICANT: Zhong, Weidong
; APPLICANT: Hong, Zhi
; APPLICANT: Ferrari, Eric
; TITLE OF INVENTION: HCV REPLICASE COMPLEXES
; FILE REFERENCE: IN01165
; CURRENT APPLICATION NUMBER: US/09/828,034
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: U.S. 60/195,852
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 36
; TYPE: RNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic RNA
US-09-828-034-5

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 36 CGCGCGCG 29

RESULT 19

US-09-733-042-31
; Sequence 31, Application US/09733042
; Patent No. US20020168709A1
; GENERAL INFORMATION:
; APPLICANT: Hennecke, Frank

APPLICANT: Renner, Wolfgang A.
TITLE OF INVENTION: Replicon Based Activation of Endogenous Genes
FILE REFERENCE: 1700.0100001
CURRENT APPLICATION NUMBER: US/09/733,042
CURRENT FILING DATE: 2000-12-11
PRIOR APPLICATION NUMBER: US 60/169,988
PRIOR FILING DATE: 1999-12-10
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn version 3.0
SEQ ID NO 31
LENGTH: 39
TYPE: DNA
ORGANISM: notins5'-FOR
US-09-733-042-31

Query Match 100.0%; Score 8; DB 9; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 2 CGCGCGCG 9

RESULT 20
US-09-733-042-31/C
Sequence 31, Application US/09733042
Patent No. US20020168709A1
GENERAL INFORMATION:
APPLICANT: Hennecke, Frank
TITLE OF INVENTION: Replicon Based Activation of Endogenous Genes
FILE REFERENCE: 1700.0100001
CURRENT APPLICATION NUMBER: US/09/733,042
CURRENT FILING DATE: 2000-12-11
PRIOR APPLICATION NUMBER: US 60/169,988
PRIOR FILING DATE: 1999-12-10
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn version 3.0
SEQ ID NO 31
LENGTH: 39
TYPE: DNA
ORGANISM: notins5'-FOR
US-09-733-042-31

Query Match 100.0%; Score 8; DB 9; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 9 CGCGCGCG 2

RESULT 21
US-10-007-132-45
Sequence 45, Application US/10007132
Publication No. US20030027254A1
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A
Borowsky, Beth
Smith, Kelli E
TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS
AND USES THEREOF
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,132
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/058,333
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52241-E/JPW/KDB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 278 0400
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-10-007-132-45

Query Match 100.0%; Score 8; DB 9; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 7 CGCGCGCG 14

RESULT 22
US-10-007-132-45/C
Sequence 45, Application US/10007132
Publication No. US20030027254A1
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A
Borowsky, Beth
Smith, Kelli E
TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS
AND USES THEREOF
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,132
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/058,333
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52241-E/JPW/KDB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 278 0400

```

; TELEFAX: 212 391 0525
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-10-007-132-45

Query Match          100.0%; Score 8; DB 9; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
    |||||||
Db 14 CGCGCGCG 7

RESULT 23
US-09-790-417-19
; Sequence 19, Application US/09790417
; Patent No. US20010031470A1
; GENERAL INFORMATION:
; APPLICANT: Shultz, John W.
; APPLICANT: Lewis, Martin K.
; APPLICANT: Lieppe, Donna
; APPLICANT: Mandrekar, Michelle
; APPLICANT: Kephart, Daniel
; APPLICANT: Rhodes, Richard B.
; APPLICANT: Andrews, Christine A.
; APPLICANT: Hartnett, James R.
; APPLICANT: Gu, Trent
; APPLICANT: Olson, Ryan J.
; APPLICANT: Wood, Keith W.
; APPLICANT: Welch, Roy
; TITLE OF INVENTION: Nucleic Acid Detection
; FILE REFERENCE: Pro-103 6868/75528
; CURRENT APPLICATION NUMBER: US/09/790,417
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 09/358,972
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/042,287
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 19
; LENGTH: 55
; TYPE: DNA
; ORGANISM: Cytomegalovirus
; FEATURE:
; OTHER INFORMATION: target for mutant cytomegalovirus
US-09-790-417-19

Query Match          100.0%; Score 8; DB 10; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
    |||||||
Db 27 CGCGCGCG 34

RESULT 24
US-09-790-417-19/c
; Sequence 19, Application US/09790417
; Patent No. US20010031470A1
; GENERAL INFORMATION:
; APPLICANT: Shultz, John W.
; APPLICANT: Lewis, Martin K.
; APPLICANT: Lieppe, Donna
; APPLICANT: Mandrekar, Michelle
```

```

; APPLICANT: Kephart, Daniel
; APPLICANT: Rhodes, Richard B.
; APPLICANT: Andrews, Christine A.
; APPLICANT: Hartnett, James R.
; APPLICANT: Gu, Trent
; APPLICANT: Olson, Ryan J.
; APPLICANT: Wood, Keith W.
; APPLICANT: Welch, Roy
; TITLE OF INVENTION: Nucleic Acid Detection
; FILE REFERENCE: Pro-103 6868/75528
; CURRENT APPLICATION NUMBER: US/09/790,417
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 09/358,972
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 09/042,287
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 19
; LENGTH: 55
; TYPE: DNA
; ORGANISM: Cytomegalovirus
; FEATURE:
; OTHER INFORMATION: target for mutant cytomegalovirus
US-09-790-417-19

Query Match          100.0%; Score 8; DB 10; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
    |||||||
Db 34 CGCGCGCG 27

RESULT 25
US-09-899-381-37
; Sequence 37, Application US/09899381
; Patent No. US20020068293A1
; GENERAL INFORMATION:
; APPLICANT: Delenstarr, Glend C.
; APPLICANT: Wolber, Paul K.
; APPLICANT: Sana, Theodore R.
; TITLE OF INVENTION: Arrays Having Background Features and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: 10010760-1
; CURRENT APPLICATION NUMBER: US/09/899,381
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/398,399
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic probe
US-09-899-381-37

Query Match          100.0%; Score 8; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
    |||||||
Db 44 CGCGCGCG 51

RESULT 26
US-09-899-381-37/c
; Sequence 37, Application US/09899381
; Patent No. US20020068293A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Delenstarr, Glend C.
; APPLICANT: Wolber, Pual K.
; APPLICANT: Sana, Theodore R.
; TITLE OF INVENTION: Arrays Having Background Features and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: 10010760-1
; CURRENT APPLICATION NUMBER: US/09/899,381
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/398,399
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic probe
US-09-899-381-37
```

```
Query Match          100.0%; Score 8; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGCGCGCG 8
    |||||
DB 17 CGCGCGCG 10
```

```
RESULT 27
US-10-014-973A-19
; Sequence 19, Application US/10014973A
; Patent No. US20020127581A1
; GENERAL INFORMATION:
; APPLICANT: Ellington, Andrew
; APPLICANT: Jhaveri, Sulay
; APPLICANT: Rajendran, Manjula
; TITLE OF INVENTION: In Vitro Selection of Signaling Aptamers
; FILE REFERENCE: D6297
; CURRENT APPLICATION NUMBER: US/10/014,973A
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/244,010
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 19
; LENGTH: 61
; TYPE: RNA
; ORGANISM: artificial sequence
; FEATURE:
; LOCATION: 48
; OTHER INFORMATION: Sequence of the cloned aptamer raf1s in
; OTHER INFORMATION: Family 1; u is 12-F-u at position 48
US-10-014-973A-19
```

```
Query Match          100.0%; Score 8; DB 12; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGCGCGCG 8
    |||||
DB 51 CGCGCGCG 58
```

```
RESULT 28
US-10-014-973A-19/c
; Sequence 19, Application US/10014973A
; Patent No. US20020127581A1
; GENERAL INFORMATION:
; APPLICANT: Ellington, Andrew
; APPLICANT: Jhaveri, Sulay
; APPLICANT: Rajendran, Manjula
; TITLE OF INVENTION: In Vitro Selection of Signaling Aptamers
```

```
; FILE REFERENCE: D6297
; CURRENT APPLICATION NUMBER: US/10/014,973A
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/244,010
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 19
; LENGTH: 61
; TYPE: RNA
; ORGANISM: artificial sequence
; FEATURE:
; LOCATION: 48
; OTHER INFORMATION: Sequence of the cloned aptamer raf1s in
; OTHER INFORMATION: Family 1; u is 12-F-u at position 48
US-10-014-973A-19
```

```
Query Match          100.0%; Score 8; DB 12; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGCGCGCG 8
    |||||
DB 58 CGCGCGCG 51
```

```
RESULT 29
US-10-057-940-8
; Sequence 8, Application US/10057940
; Patent No. US20020168686A1
; GENERAL INFORMATION:
; APPLICANT: Pantoliaro, Michael W.
; APPLICANT: Saleme, F. Raymond
; APPLICANT: Carver, Jr., Theodore, E.
; TITLE OF INVENTION: High Throughput Method for Functionally Classifying Proteins
; TITLE OF INVENTION: Identified using a Genomics Approach
; FILE REFERENCE: 1503.0310002/7AG/7SO
; CURRENT APPLICATION NUMBER: US/10/057,940
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 09/190,128
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/065,129
; PRIOR FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 64
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GC-rich tract
US-10-057-940-8
```

```
Query Match          100.0%; Score 8; DB 9; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGCGCGCG 8
    |||||
DB 55 CGCGCGCG 62
```

```
RESULT 30
US-10-057-940-8/c
; Sequence 8, Application US/10057940
; Patent No. US20020168686A1
; GENERAL INFORMATION:
; APPLICANT: Pantoliaro, Michael W.
; APPLICANT: Saleme, F. Raymond
; APPLICANT: Carver, Jr., Theodore, E.
; TITLE OF INVENTION: High Throughput Method for Functionally Classifying Proteins
; TITLE OF INVENTION: Identified using a Genomics Approach
; FILE REFERENCE: 1503.0310002/7AG/7SO
; CURRENT APPLICATION NUMBER: US/10/057,940
```

```

; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 09/190,128
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/065,129
; PRIOR FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 64
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GC-rich tract
US-10-057-940-8

Query Match
Best Local Similarity 100.0%; Score 8; DB 9; Length 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 9 CGCGCGCG 2

RESULT 31
US-09-473-872-54
; Sequence 54, Application US/09473872
; Patent No. US20020064876A1
; GENERAL INFORMATION:
; APPLICANT: Yoon, Kyonggeun
; TITLE OF INVENTION: NO. US20020064876A1 Gene Therapy Methods For The Treatment of
; FILE REFERENCE: Disorders
; CURRENT APPLICATION NUMBER: US/09/473,872
; CURRENT FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 54
; LENGTH: 68
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule:
; OTHER INFORMATION: Chimeric oligonucleotide
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
US-09-473-872-54

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 52 CGCGCGCG 59

RESULT 32
US-09-473-872-54/c
; Sequence 54, Application US/09473872
; Patent No. US20020064876A1
; GENERAL INFORMATION:
; APPLICANT: Yoon, Kyonggeun
; TITLE OF INVENTION: NO. US20020064876A1 Gene Therapy Methods For The Treatment of
; FILE REFERENCE: Disorders
; CURRENT APPLICATION NUMBER: US/09/473,872
; CURRENT FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 54
; LENGTH: 68
; TYPE: DNA
```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule:
; OTHER INFORMATION: Chimeric oligonucleotide
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
US-09-473-872-54

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 59 CGCGCGCG 52

RESULT 33
US-09-835-976B-122
; Sequence 122, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 122
; LENGTH: 72
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-122

Query Match
Best Local Similarity 100.0%; Score 8; DB 9; Length 72;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 30 CGCGCGCG 37

RESULT 34
US-09-835-976B-122/c
; Sequence 122, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 122
; LENGTH: 72
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-122

Query Match
Best Local Similarity 100.0%; Score 8; DB 9; Length 72;
```

Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|11111111|
Db 37 CGCGCGCG 30

RESULT 35

US-09-835-976B-119
; Sequence 119, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: ATTORNEY DOCKET NO. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 74
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-119

Query Match 100.0%; Score 8; DB 9; Length 74;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|11111111|
Db 34 CGCGCGCG 41

RESULT 36

US-09-835-976B-119/c
; Sequence 119, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: ATTORNEY DOCKET NO. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 74
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-119

Query Match 100.0%; Score 8; DB 9; Length 74;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|11111111|
Db 41 CGCGCGCG 34

RESULT 37

US-09-835-976B-127
; Sequence 127, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: ATTORNEY DOCKET NO. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 127
; LENGTH: 76
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-127

Query Match 100.0%; Score 8; DB 9; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|11111111|
Db 32 CGCGCGCG 39

RESULT 38

US-09-835-976B-127/c
; Sequence 127, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: ATTORNEY DOCKET NO. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 127
; LENGTH: 76
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-127

Query Match 100.0%; Score 8; DB 9; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
|11111111|
Db 41 CGCGCGCG 34

RESULT 39

US-09-835-976B-129
; Sequence 129, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC

```
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 129
; LENGTH: 76
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-129

Query Match          100.0%; Score 8; DB 9; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 34 CGCGCGCG 41

RESULT 40
US-09-835-976B-129/C
; Sequence 129, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 129
; LENGTH: 76
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-129

Query Match          100.0%; Score 8; DB 9; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 41 CGCGCGCG 34

RESULT 41
US-09-835-976B-126
; Sequence 126, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 126
; LENGTH: 78

; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-126

Query Match          100.0%; Score 8; DB 9; Length 78;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 43 CGCGCGCG 36

RESULT 42
US-09-835-976B-126/C
; Sequence 126, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC
; TITLE OF INVENTION: POLYPEPTIDES AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 126
; LENGTH: 78
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-126

Query Match          100.0%; Score 8; DB 9; Length 78;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 43 CGCGCGCG 36

RESULT 43
US-09-835-976B-128
; Sequence 128, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delpire, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: Attorney Docket No. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 128
; LENGTH: 78
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-128

Query Match          100.0%; Score 8; DB 9; Length 78;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
```

Db 34 CGCGCGCG 41
|||||||

RESULT 44
US-09-835-976B-128/c
; Sequence 128, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delplre, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: ATTORNEY DOCKET NO. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 128
; LENGTH: 78
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-128

Query Match 100.0%; Score 8; DB 9; Length 78;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||

Db 43 CGCGCGCG 36

RESULT 45
US-09-835-976B-124
; Sequence 124, Application US/09835976B
; Publication No. US20030027983A1
; GENERAL INFORMATION:
; APPLICANT: Mount, David B.
; APPLICANT: Delplre, Eric
; APPLICANT: Gamba, Gerardo
; APPLICANT: Alfred L. George, Jr.
; TITLE OF INVENTION: PURIFIED AND ISOLATED POTASSIUM-CHLORIDE COTRANSPORTER NUCLEIC AC
; TITLE OF INVENTION: POLYPEPTIDES AND
; FILE REFERENCE: ATTORNEY DOCKET NO. US20030027983A1 1242-26-2
; CURRENT APPLICATION NUMBER: US/09/835,976B
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 124
; LENGTH: 80
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-835-976B-124

Query Match 100.0%; Score 8; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||

Db 36 CGCGCGCG 43

Search completed: March 14, 2003, 04:34:58
Job time : 65 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 03:37:53 ; Search time 43 Seconds
(without alignments)
57.056 Million cell updates/sec

Title: CGCGCGCG

Perfect score: 8

Sequence: 1 cgcgcgcg 8

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 44362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/pdata/1/lna/5A.COMB.seq: *
2: /cgn2_6/pdata/1/lna/5B.COMB.seq: *
3: /cgn2_6/pdata/1/lna/6A.COMB.seq: *
4: /cgn2_6/pdata/1/lna/6B.COMB.seq: *
5: /cgn2_6/pdata/1/lna/PCRTUS.COMB.seq: *
6: /cgn2_6/pdata/1/lna/Backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	100.0	10	2	US-08-734-973-9
2	8	100.0	10	2	US-08-734-973-9
3	8	100.0	10	2	US-08-734-973-10
4	8	100.0	10	2	US-08-734-973-10
5	8	100.0	10	2	US-08-734-973-11
6	8	100.0	10	2	US-08-734-973-11
7	8	100.0	10	2	US-08-734-973-12
8	8	100.0	10	2	US-08-734-973-12
9	8	100.0	10	2	US-08-734-973-13
10	8	100.0	10	2	US-08-734-973-13
11	8	100.0	10	2	US-08-734-973-14
12	8	100.0	10	2	US-08-734-973-14
13	8	100.0	10	2	US-08-734-973-15
14	8	100.0	10	2	US-08-734-973-15
15	8	100.0	10	2	US-08-734-973-16
16	8	100.0	10	2	US-08-734-973-16
17	8	100.0	10	2	US-08-734-973-17
18	8	100.0	10	2	US-08-734-973-17
19	8	100.0	10	2	US-08-734-973-18
20	8	100.0	10	2	US-08-734-973-18
21	8	100.0	10	2	US-08-734-973-19
22	8	100.0	10	2	US-08-734-973-19
23	8	100.0	10	2	US-08-734-973-20
24	8	100.0	10	2	US-08-734-973-20
25	8	100.0	10	2	US-08-734-973-21
26	8	100.0	10	2	US-08-734-973-21
27	8	100.0	10	2	US-08-734-973-22

c 28	8	100.0	10	2	US-08-734-973-22	Sequence 22, Appl
c 29	8	100.0	10	2	US-08-734-973-23	Sequence 23, Appl
c 30	8	100.0	10	2	US-08-734-973-23	Sequence 23, Appl
c 31	8	100.0	10	2	US-08-734-973-24	Sequence 24, Appl
c 32	8	100.0	10	2	US-08-734-973-24	Sequence 24, Appl
c 33	8	100.0	10	2	US-08-734-973-25	Sequence 25, Appl
c 34	8	100.0	10	2	US-08-734-973-25	Sequence 25, Appl
c 35	8	100.0	10	2	US-08-734-973-26	Sequence 26, Appl
c 36	8	100.0	10	2	US-08-734-973-26	Sequence 26, Appl
c 37	8	100.0	10	2	US-08-734-973-27	Sequence 27, Appl
c 38	8	100.0	10	2	US-08-734-973-27	Sequence 27, Appl
c 39	8	100.0	10	3	US-08-729-598-13	Sequence 13, Appl
c 40	8	100.0	10	3	US-08-729-598-13	Sequence 13, Appl
c 41	8	100.0	12	4	US-09-393-783A-79	Sequence 79, Appl
c 42	8	100.0	12	4	US-09-393-783A-79	Sequence 79, Appl
c 43	8	100.0	12	4	US-09-151-880B-79	Sequence 79, Appl
c 44	8	100.0	12	4	US-09-151-880B-79	Sequence 79, Appl
c 45	8	100.0	14	2	US-08-595-043A-44	Sequence 44, Appl

ALIGNMENTS

RESULT 1
US-08-734-973-9
; Sequence 9, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoller, Daniel L.
; APPLICANT: Basik, Mark
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 9 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHEICAL: NO
; US-08-734-973-9

Query Match 100.0%; Score 8; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.2e+04; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

```
RESULT 2
US-08-734-973-9/c
; Sequence 9, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 849-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 9 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; US-08-734-973-9

Query Match          100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 3
US-08-734-973-10
; Sequence 10, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 849-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 10 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
```

```
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 849-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 10 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO

US-08-734-973-10

Query Match          100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8
```

```
RESULT 4
US-08-734-973-10/c
; Sequence 10, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 849-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 10 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
```

HYPOTHETICAL: NO

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 5

US-08-734-973-11
Sequence 11, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Met Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 Inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single-stranded

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-734-973-11

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 6

US-08-734-973-11/C

Sequence 11, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Met Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 Inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TITLE OF INVENTION: Genomic Instability

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Met Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 Inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single-stranded

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-734-973-11

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 7

US-08-734-973-12

Sequence 12, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Met Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 Inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 12 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-12

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
DB 1 CGCGCGCG 8

RESULT 8
US-08-734-973-12/C
; Sequence 12, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 12 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
US-08-734-973-12

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
DB 8 CGCGCGCG 1

RESULT 9
US-08-734-973-13
; Sequence 13, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 13 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
US-08-734-973-13

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
DB 1 CGCGCGCG 8

RESULT 10
US-08-734-973-13/C
; Sequence 13, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch

```
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 13 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-13
```

```
Query Match          100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 CGCGCGCG 8
Db 8 CGCGCGCG 1
```

```
RESULT 11
US-08-734-973-14
Sequence 14, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 14 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
```

```
US-08-734-973-14
```

```
Query Match          100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 CGCGCGCG 8
Db 1 CGCGCGCG 8
```

```
RESULT 12
US-08-734-973-14/C
Sequence 14, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 14 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-14

Query Match          100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 13
US-08-734-973-15
Sequence 15, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
```

NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 15 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-15

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 14
US-08-734-973-15/c
Sequence 15 Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
APPLICANT: Anderson, Garth R.
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 15 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-15

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 15
US-08-734-973-16
Sequence 16, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
APPLICANT: Anderson, Garth R.
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 16 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-16

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

```
RESULT 16
US-08-734-973-16/c
; Sequence 16, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 16 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; US-08-734-973-16

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 17
US-08-734-973-17
; Sequence 17, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 17 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; US-08-734-973-17
```

```
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 17 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-17

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 18
US-08-734-973-17/c
; Sequence 17, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; TITLE OF INVENTION: Genomic Instability
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 17 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; US-08-734-973-17
```

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 19

US-08-734-973-18
Sequence 18, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Mt Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 18 :

SEQUENCE CHARACTERISTICS:

LENGTH: 10 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single-stranded

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-734-973-18

Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 10;
Pred. No. 1.2e+04;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 20

US-08-734-973-18/c

Sequence 18, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Mt Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 18 :

SEQUENCE CHARACTERISTICS:

LENGTH: 10 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single-stranded

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-734-973-18

Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 10;
Pred. No. 1.2e+04;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 21

US-08-734-973-19

Sequence 19, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Mt Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 19
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-19

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 22

US-08-734-973-19/c
Sequence 19, Application US/08734973
Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

TITLE OF INVENTION: Genomic Instability

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Met Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 Inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 19

SEQUENCE CHARACTERISTICS:

LENGTH: 10 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single-stranded

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-734-973-19

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 23

US-08-734-973-20

Sequence 20, Application US/08734973
Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

TITLE OF INVENTION: Genomic Instability

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Met Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 Inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 20

SEQUENCE CHARACTERISTICS:

LENGTH: 10 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single-stranded

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-734-973-20

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 24

US-08-734-973-20/c
Sequence 20, Application US/08734973
Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

TITLE OF INVENTION: Genomic Instability

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Met Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 Inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 20 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-20

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
DB 8 CGCGCGCG 1

RESULT 25
US-08-734-973-21
Sequence 21, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 21 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-21

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
DB 1 CGCGCGCG 8

RESULT 26
US-08-734-973-21/c
Sequence 21, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 21 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-21

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
|||||||
DB 8 CGCGCGCG 1

RESULT 27
US-08-734-973-22
Sequence 22, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 22 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; US-08-734-973-22

Query Match          100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 28
US-08-734-973-22/c
; Sequence 22, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349

```

```

; INFORMATION FOR SEQ ID NO: 22 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; US-08-734-973-22

Query Match          100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 29
US-08-734-973-23
; Sequence 23, Application US/08734973
; Patent No. 5912147
; GENERAL INFORMATION:
; APPLICANT: Stoler, Daniel L.
; APPLICANT: Basik, Mark
; APPLICANT: Anderson, Garth R.
; TITLE OF INVENTION: A Rapid Means For Quantitating
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,973
; FILING DATE: October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 03551.0021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 23 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; US-08-734-973-23

Query Match          100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 30
US-08-734-973-23/c

```

Sequence 23, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
APPLICANT: Anderson, Garth R.
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 23 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-23

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 31
US-08-734-973-24
Sequence 24, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
APPLICANT: Anderson, Garth R.
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 24 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-24

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 32
US-08-734-973-24/C
Sequence 24, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
APPLICANT: Anderson, Garth R.
TITLE OF INVENTION: A Rapid Means For Quantitating
TITLE OF INVENTION: Genomic Instability
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 24 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-24

Query Match 100.0%; Score 8; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
11111111
DB 8 CGCGCGCG 1

RESULT 33

US-08-734-973-25

Sequence 25, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Mt Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 25 :

SEQUENCE CHARACTERISTICS:

LENGTH: 10 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single-stranded

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-734-973-25

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
11111111
DB 1 CGCGCGCG 8

RESULT 34

US-08-734-973-25/c

Sequence 25, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Mt Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 25 :

SEQUENCE CHARACTERISTICS:

LENGTH: 10 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single-stranded

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-734-973-25

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
11111111
DB 8 CGCGCGCG 1

RESULT 35

US-08-734-973-26

Sequence 26, Application US/08734973

Patent No. 5912147

GENERAL INFORMATION:

APPLICANT: Stoler, Daniel L.

APPLICANT: Basik, Mark

APPLICANT: Anderson, Garth R.

TITLE OF INVENTION: A Rapid Means For Quantitating

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear

STREET: 1800 One Mt Plaza

CITY: Buffalo

STATE: New York

COUNTRY: United States

ZIP: 14203-2391

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS/ Microsoft Windows

SOFTWARE: Wordperfect for Windows

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973

FILING DATE: October 1996

ATTORNEY/AGENT INFORMATION:

NAME: Nelson, M. Bud

REGISTRATION NUMBER: 35,300

REFERENCE/DOCKET NUMBER: 03551.0021

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 856-4000

TELEFAX: (716) 849-0349

INFORMATION FOR SEQ ID NO: 26 :

SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-26

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 36
US-08-734-973-26/C
Sequence 26, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: A Rapid Means For Quantitating
NUMBER OF INVENTION: Genomic Instability
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 26 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-26

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 37
US-08-734-973-27
Sequence 27, Application US/08734973

Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: A Rapid Means For Quantitating
NUMBER OF INVENTION: Genomic Instability
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 27 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-27

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 38
US-08-734-973-27/C
Sequence 27, Application US/08734973
Patent No. 5912147
GENERAL INFORMATION:
APPLICANT: Stoler, Daniel L.
APPLICANT: Basik, Mark
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: A Rapid Means For Quantitating
NUMBER OF INVENTION: Genomic Instability
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Mt Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect for Windows
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/734,973
FILING DATE: October 1996
ATTORNEY/AGENT INFORMATION:
NAME: Nelson, M. Bud
REGISTRATION NUMBER: 35,300
REFERENCE/DOCKET NUMBER: 03551.0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 27 :
SEQUENCE CHARACTERISTICS:
LENGTH: 10 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-08-734-973-27

Query Match 100.0%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 39

US-08-729-598-13
Sequence 13, Application US/08729598
Patent No. 6001657
GENERAL INFORMATION:
APPLICANT: Hardin, Charles C.
APPLICANT: Brown II, Bernard A.
APPLICANT: Roberts, John J.
APPLICANT: Peltue, Stephen A.
TITLE OF INVENTION: Antibodies That Selectively Bind
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sorojini J. Biswas
STREET: P.O. Box 37428
CITY: Raleigh
STATE: No. 6001657th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,598
FILING DATE: 11-OCT-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Biswas, Sorojini J.
REGISTRATION NUMBER: 39,111
REFERENCE/DOCKET NUMBER: 5051-301A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 854-1400
TELEFAX: (919) 854-1401
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: DNA (genomic)
US-08-729-598-13

Query Match 100.0%; Score 8; DB 3; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 40

US-08-729-598-13/c
Sequence 13, Application US/08729598
Patent No. 6001657
GENERAL INFORMATION:
APPLICANT: Hardin, Charles C.
APPLICANT: Brown II, Bernard A.
APPLICANT: Roberts, John J.
APPLICANT: Peltue, Stephen A.
TITLE OF INVENTION: Antibodies That Selectively Bind
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sorojini J. Biswas
STREET: P.O. Box 37428
CITY: Raleigh
STATE: No. 6001657th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,598
FILING DATE: 11-OCT-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Biswas, Sorojini J.
REGISTRATION NUMBER: 39,111
REFERENCE/DOCKET NUMBER: 5051-301A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 854-1400
TELEFAX: (919) 854-1401
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: DNA (genomic)
US-08-729-598-13

Query Match 100.0%; Score 8; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
DB 10 CGCGCGCG 3

RESULT 41

US-09-393-783A-79
Sequence 79, Application US/09393783A
Patent No. 6355428
GENERAL INFORMATION:
APPLICANT: Schroth, Gary P.
APPLICANT: Bruice, Thomas Wayne
APPLICANT: Sub. Young J.
TITLE OF INVENTION: Nucleic Acid Ligand Interaction Assays
FILE REFERENCE: 4600-0128.30
CURRENT APPLICATION NUMBER: US/09/393,783A
CURRENT FILING DATE: 1999-10-09

```

; PRIOR APPLICATION NUMBER: US 09/151,890
; PRIOR FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (1)...(12)
; OTHER INFORMATION: synthesized test oligonucleotide for binding
; OTHER INFORMATION: studies
; US-09-393-783A-79

```

```

Query Match          100.0%; Score 8; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CGCGCGCG 8
    |||||
Db 3 CGCGCGCG 10

```

```

RESULT 42
; US-09-393-783A-79/c
; Sequence 79, Application US/09393783A
; Patent No. 6353428
; GENERAL INFORMATION:
; APPLICANT: Schroth, Gary P.
; APPLICANT: Bruice, Thomas Wayne
; APPLICANT: Suh, Young J.
; TITLE OF INVENTION: Nucleic Acid Ligand Interaction Assays
; FILE REFERENCE: 4600-0128.30
; CURRENT APPLICATION NUMBER: US/09/393,783A
; PRIOR FILING DATE: 1999-10-09
; PRIOR APPLICATION NUMBER: US 09/151,890
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (1)...(12)
; OTHER INFORMATION: synthesized test oligonucleotide for binding
; OTHER INFORMATION: studies
; US-09-393-783A-79

```

```

Query Match          100.0%; Score 8; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CGCGCGCG 8
    |||||
Db 10 CGCGCGCG 3

```

```

RESULT 43
; US-09-151-890B-79
; Sequence 79, Application US/09151890B
; Patent No. 6420109
; GENERAL INFORMATION:
; APPLICANT: Gary P. Schroth
; APPLICANT: Thomas Wayne Bruice
; APPLICANT: Young J. Suh
; TITLE OF INVENTION: Nucleic Acid Ligand Interaction Assays
; FILE REFERENCE: 4600-0128
; CURRENT APPLICATION NUMBER: US/09/151,890B
; CURRENT FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 80

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (1)...(12)
; OTHER INFORMATION: synthesized test oligonucleotide for binding
; OTHER INFORMATION: studies
; US-09-151-890B-79

```

```

Query Match          100.0%; Score 8; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CGCGCGCG 8
    |||||
Db 3 CGCGCGCG 10

```

```

RESULT 44
; US-09-151-890B-79/c
; Sequence 79, Application US/09151890B
; Patent No. 6420109
; GENERAL INFORMATION:
; APPLICANT: Gary P. Schroth
; APPLICANT: Thomas Wayne Bruice
; APPLICANT: Young J. Suh
; TITLE OF INVENTION: Nucleic Acid Ligand Interaction Assays
; FILE REFERENCE: 4600-0128
; CURRENT APPLICATION NUMBER: US/09/151,890B
; CURRENT FILING DATE: 1998-09-11
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (1)...(12)
; OTHER INFORMATION: synthesized test oligonucleotide for binding
; OTHER INFORMATION: studies
; US-09-151-890B-79

```

```

Query Match          100.0%; Score 8; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CGCGCGCG 8
    |||||
Db 10 CGCGCGCG 3

```

```

RESULT 45
; US-08-595-043A-44
; Sequence 44, Application US/08595043A
; Patent No. 5935824
; GENERAL INFORMATION:
; APPLICANT: SGARIATO, GREGORY D.
; TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/595,043A
; FILING DATE: 31-JAN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: SGAR-00371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-595-043A-44

```

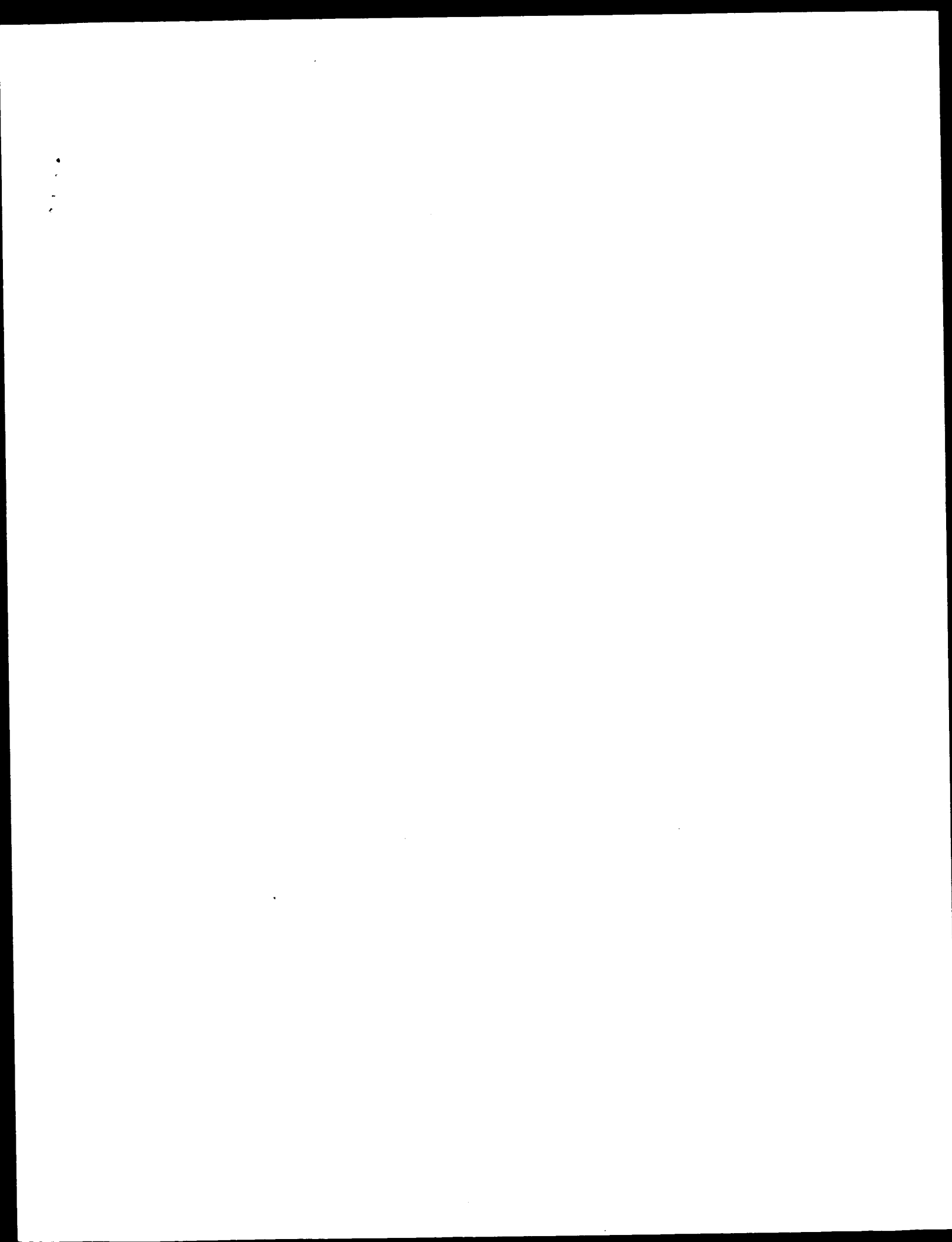
```

Query Match      100.0%; Score 8; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
   |||||
Db 6 CGCGCGCG 13

```

Search completed: March 14, 2003, 04:33:32
 Job time : 44 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 02:36:13 ; Search time 1363 Seconds
(without alignments)
170.816 Million cell updates/sec

Title: CGCGCGCG

Perfect score: 8

Sequence: 1 cgcgcgcg 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

Geneml: *
1: gb_da: *
2: gb_hcg: *
3: gb_in: *
4: gb_com: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_scs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rpd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_ay: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	100.0	10	AR071780	Sequence
2	8	100.0	10	AR071780	Sequence
3	8	100.0	10	AR071781	Sequence
4	8	100.0	10	AR071781	Sequence
5	8	100.0	10	AR071782	Sequence
6	8	100.0	10	AR071782	Sequence
7	8	100.0	10	AR071783	Sequence
8	8	100.0	10	AR071783	Sequence
9	8	100.0	10	AR071784	Sequence
10	8	100.0	10	AR071784	Sequence
11	8	100.0	10	AR071785	Sequence
12	8	100.0	10	AR071785	Sequence
13	8	100.0	10	AR071786	Sequence
14	8	100.0	10	AR071786	Sequence
15	8	100.0	10	AR071787	Sequence
16	8	100.0	10	AR071787	Sequence
17	8	100.0	10	AR071788	Sequence
18	8	100.0	10	AR071788	Sequence
19	8	100.0	10	AR071789	Sequence
20	8	100.0	10	AR071789	Sequence
21	8	100.0	10	AR071790	Sequence
22	8	100.0	10	AR071790	Sequence
23	8	100.0	10	AR071791	Sequence
24	8	100.0	10	AR071791	Sequence
25	8	100.0	10	AR071792	Sequence
26	8	100.0	10	AR071792	Sequence
27	8	100.0	10	AR071793	Sequence
28	8	100.0	10	AR071793	Sequence
29	8	100.0	10	AR071794	Sequence
30	8	100.0	10	AR071794	Sequence
31	8	100.0	10	AR071795	Sequence
32	8	100.0	10	AR071795	Sequence
33	8	100.0	10	AR071796	Sequence
34	8	100.0	10	AR071796	Sequence
35	8	100.0	10	AR071797	Sequence
36	8	100.0	10	AR071797	Sequence
37	8	100.0	10	AR071798	Sequence
38	8	100.0	10	AR071798	Sequence
39	8	100.0	10	AR071798	Sequence
40	8	100.0	10	AR094565	Sequence
41	8	100.0	12	AR193370	Sequence
42	8	100.0	12	AR193370	Sequence
43	8	100.0	16	AR048207	Sequence
44	8	100.0	16	AR048207	Sequence
45	8	100.0	16	AR055680	Sequence

ALIGNMENTS

RESULT 1
LOCUS AR071780 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 9 from patent US 5912147.
ACCESSION AR071780
VERSION AR071780.1 GI:7222668
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 9 15-JUN-1999;
FEATURES Location/Qualifiers

```

source 1..10
/organism="unknown"
BASE COUNT 1 a 4 c 5 g 0 t
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 2
AR071780/c AR071780 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 9 from patent US 5912147.
ACCESSION AR071780
VERSION AR071780.1 GI:7222668
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 10)
TITLE Stoler,D., Basik,M. and Anderson,G.
JOURNAL Patent: US 5912147-A 9 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"
BASE COUNT 1 a 4 c 5 g 0 t
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 3
AR071781 AR071781 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 10 from patent US 5912147.
ACCESSION AR071781
VERSION AR071781.1 GI:7222669
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 10)
TITLE Stoler,D., Basik,M. and Anderson,G.
JOURNAL Patent: US 5912147-A 10 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"
BASE COUNT 0 a 4 c 6 g 0 t
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 4

```

```

AR071781/c AR071781 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 10 from patent US 5912147.
ACCESSION AR071781
VERSION AR071781.1 GI:7222669
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 10)
TITLE Stoler,D., Basik,M. and Anderson,G.
JOURNAL Patent: US 5912147-A 10 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"
BASE COUNT 0 a 4 c 6 g 0 t
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 5
AR071782 AR071782 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 11 from patent US 5912147.
ACCESSION AR071782
VERSION AR071782.1 GI:7222670
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 10)
TITLE Stoler,D., Basik,M. and Anderson,G.
JOURNAL Patent: US 5912147-A 11 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"
BASE COUNT 1 a 5 c 4 g 0 t
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 6
AR071782/c AR071782 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 11 from patent US 5912147.
ACCESSION AR071782
VERSION AR071782.1 GI:7222670
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 10)
TITLE Stoler,D., Basik,M. and Anderson,G.
JOURNAL Patent: US 5912147-A 11 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10

```

```

BASE COUNT      1 a      5 c      4 g      0 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 8; DB 6; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
    |||||
Db 8 CGCGCGCG 1

RESULT 7
AR071783
LOCUS      AR071783      10 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 12 from patent US 5912147.
ACCESSION  AR071783
VERSION     AR071783.1 GI:7222671
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 10)
AUTHORS     Stoler,D., Basik,M. and Anderson,G.
TITLE       Rapid means of quantitating genomic instability
JOURNAL     Patent: US 5912147-A 12 15-JUN-1999;
FEATURES
    source 1..10
            /organism="unknown"

BASE COUNT      1 a      4 c      4 g      1 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 8; DB 6; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
    |||||
Db 1 CGCGCGCG 8

RESULT 8
AR071783/c
LOCUS      AR071783      10 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 12 from patent US 5912147.
ACCESSION  AR071783
VERSION     AR071783.1 GI:7222671
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 10)
AUTHORS     Stoler,D., Basik,M. and Anderson,G.
TITLE       Rapid means of quantitating genomic instability
JOURNAL     Patent: US 5912147-A 12 15-JUN-1999;
FEATURES
    source 1..10
            /organism="unknown"

BASE COUNT      1 a      4 c      4 g      1 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 8; DB 6; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
    |||||
Db 8 CGCGCGCG 1

RESULT 9
AR071784
LOCUS      AR071784      10 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 13 from patent US 5912147.
ACCESSION  AR071784
VERSION     AR071784.1 GI:7222672
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 10)
AUTHORS     Stoler,D., Basik,M. and Anderson,G.
TITLE       Rapid means of quantitating genomic instability
JOURNAL     Patent: US 5912147-A 13 15-JUN-1999;
FEATURES
    source 1..10
            /organism="unknown"

BASE COUNT      1 a      4 c      4 g      1 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 8; DB 6; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
    |||||
Db 1 CGCGCGCG 8

RESULT 10
AR071784/c
LOCUS      AR071784      10 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 13 from patent US 5912147.
ACCESSION  AR071784
VERSION     AR071784.1 GI:7222672
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 10)
AUTHORS     Stoler,D., Basik,M. and Anderson,G.
TITLE       Rapid means of quantitating genomic instability
JOURNAL     Patent: US 5912147-A 13 15-JUN-1999;
FEATURES
    source 1..10
            /organism="unknown"

BASE COUNT      1 a      4 c      4 g      1 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 8; DB 6; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
    |||||
Db 8 CGCGCGCG 1

RESULT 11
AR071785
LOCUS      AR071785      10 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 14 from patent US 5912147.
ACCESSION  AR071785
VERSION     AR071785.1 GI:7222673
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 10)
AUTHORS     Stoler,D., Basik,M. and Anderson,G.
TITLE       Rapid means of quantitating genomic instability
JOURNAL     Patent: US 5912147-A 14 15-JUN-1999;
FEATURES
    source 1..10
            /organism="unknown"

```

```

LOCUS      AR071784      10 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 13 from patent US 5912147.
ACCESSION  AR071784
VERSION     AR071784.1 GI:7222672
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 10)
AUTHORS     Stoler,D., Basik,M. and Anderson,G.
TITLE       Rapid means of quantitating genomic instability
JOURNAL     Patent: US 5912147-A 13 15-JUN-1999;
FEATURES
    source 1..10
            /organism="unknown"

BASE COUNT      1 a      4 c      4 g      1 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 8; DB 6; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
    |||||
Db 1 CGCGCGCG 8

RESULT 10
AR071784/c
LOCUS      AR071784      10 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 13 from patent US 5912147.
ACCESSION  AR071784
VERSION     AR071784.1 GI:7222672
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 10)
AUTHORS     Stoler,D., Basik,M. and Anderson,G.
TITLE       Rapid means of quantitating genomic instability
JOURNAL     Patent: US 5912147-A 13 15-JUN-1999;
FEATURES
    source 1..10
            /organism="unknown"

BASE COUNT      1 a      4 c      4 g      1 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 8; DB 6; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
    |||||
Db 8 CGCGCGCG 1

RESULT 11
AR071785
LOCUS      AR071785      10 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION Sequence 14 from patent US 5912147.
ACCESSION  AR071785
VERSION     AR071785.1 GI:7222673
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 10)
AUTHORS     Stoler,D., Basik,M. and Anderson,G.
TITLE       Rapid means of quantitating genomic instability
JOURNAL     Patent: US 5912147-A 14 15-JUN-1999;
FEATURES
    source 1..10
            /organism="unknown"

```

BASE COUNT 0 a 5 c 5 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 12

AR071785/c AR071785 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 14 from patent US 5912147.
ACCESSION AR071785
VERSION AR071785.1 GI:7222673
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 14 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10

BASE COUNT 0 a 5 c 5 g 0 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 13

AR071786 AR071786 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 15 from patent US 5912147.
ACCESSION AR071786
VERSION AR071786.1 GI:7222674
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 15 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10

BASE COUNT 0 a 4 c 5 g 1 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 14

AR071786/c AR071786 10 bp DNA linear PAT 18-FEB-2000
LOCUS

DEFINITION Sequence 15 from patent US 5912147.
ACCESSION AR071786
VERSION AR071786.1 GI:7222674
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 15 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10

BASE COUNT 0 a 4 c 5 g 1 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 15

AR071787 AR071787 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 16 from patent US 5912147.
ACCESSION AR071787
VERSION AR071787.1 GI:7222675
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 16 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10

BASE COUNT 0 a 4 c 5 g 1 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 16

AR071787/c AR071787 10 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 16 from patent US 5912147.
ACCESSION AR071787
VERSION AR071787.1 GI:7222675
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 16 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10

BASE COUNT 0 a 4 c 5 g 1 t
ORIGIN

ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 17
AR071788
LOCUS AR071788 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 17 from patent US 5912147.
ACCESSION AR071788
VERSION AR071788.1 GI:7222676
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 17 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 2 a 4 c 4 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 18
AR071788/c
LOCUS AR071788 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 17 from patent US 5912147.
ACCESSION AR071788
VERSION AR071788.1 GI:7222676
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 17 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 2 a 4 c 4 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 19
AR071789
LOCUS AR071789 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 18 from patent US 5912147.

ACCESSION AR071789
VERSION AR071789.1 GI:7222677
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 18 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 1 a 4 c 5 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 20
AR071789/c
LOCUS AR071789 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 18 from patent US 5912147.
ACCESSION AR071789
VERSION AR071789.1 GI:7222677
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 18 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 1 a 4 c 5 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 21
AR071790
LOCUS AR071790 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 19 from patent US 5912147.
ACCESSION AR071790
VERSION AR071790.1 GI:7222678
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 19 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 0 a 6 c 4 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||||
 DB 1 CGCGCGCG 8

RESULT 22

AR071790/c AR071790 10 bp DNA linear PAT 18-FEB-2000
 LOCUS Sequence 19 from patent US 5912147.
 ACCESSION AR071790
 VERSION AR071790.1 GI:7222678
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE

1 (bases 1 to 10)
 AUTHORS Stoler,D., Basik,M., and Anderson,G.
 TITLE Rapid means of quantitating genomic instability
 JOURNAL Patent: US 5912147-A 19 15-JUN-1999;
 FEATURES Location/Qualifiers
 source 1..10
 /organism="unknown"

BASE COUNT 0 a 6 c 4 g 0 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||||
 DB 8 CGCGCGCG 1

RESULT 23

AR071791 AR071791 10 bp DNA linear PAT 18-FEB-2000
 LOCUS Sequence 20 from patent US 5912147.
 ACCESSION AR071791
 VERSION AR071791.1 GI:7222679
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE

1 (bases 1 to 10)
 AUTHORS Stoler,D., Basik,M., and Anderson,G.
 TITLE Rapid means of quantitating genomic instability
 JOURNAL Patent: US 5912147-A 20 15-JUN-1999;
 FEATURES Location/Qualifiers
 source 1..10
 /organism="unknown"

BASE COUNT 0 a 5 c 4 g 1 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||||
 DB 1 CGCGCGCG 8

RESULT 24

AR071791/c AR071791 10 bp DNA linear PAT 18-FEB-2000
 LOCUS Sequence 20 from patent US 5912147.
 ACCESSION AR071791

VERSION AR071791.1 GI:7222679

KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE

1 (bases 1 to 10)
 AUTHORS Stoler,D., Basik,M., and Anderson,G.
 TITLE Rapid means of quantitating genomic instability
 JOURNAL Patent: US 5912147-A 20 15-JUN-1999;
 FEATURES Location/Qualifiers
 source 1..10
 /organism="unknown"

BASE COUNT 0 a 5 c 4 g 1 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||||
 DB 8 CGCGCGCG 1

RESULT 25

AR071792 AR071792 10 bp DNA linear PAT 18-FEB-2000
 LOCUS Sequence 21 from patent US 5912147.
 ACCESSION AR071792
 VERSION AR071792.1 GI:7222680
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE

1 (bases 1 to 10)
 AUTHORS Stoler,D., Basik,M., and Anderson,G.
 TITLE Rapid means of quantitating genomic instability
 JOURNAL Patent: US 5912147-A 21 15-JUN-1999;
 FEATURES Location/Qualifiers
 source 1..10
 /organism="unknown"

BASE COUNT 0 a 5 c 4 g 1 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||||
 DB 1 CGCGCGCG 8

RESULT 26

AR071792/c AR071792 10 bp DNA linear PAT 18-FEB-2000
 LOCUS Sequence 21 from patent US 5912147.
 ACCESSION AR071792
 VERSION AR071792.1 GI:7222680
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE

1 (bases 1 to 10)
 AUTHORS Stoler,D., Basik,M., and Anderson,G.
 TITLE Rapid means of quantitating genomic instability
 JOURNAL Patent: US 5912147-A 21 15-JUN-1999;
 FEATURES Location/Qualifiers
 source 1..10
 /organism="unknown"

BASE COUNT 0 a 5 c 4 g 1 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 27
AR071793
LOCUS AR071793 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 22 from patent US 5912147.
ACCESSION AR071793
VERSION AR071793.1 GI:7222681
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 22 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 0 a 4 c 4 g 2 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 28
AR071793/c
LOCUS AR071793 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 22 from patent US 5912147.
ACCESSION AR071793
VERSION AR071793.1 GI:7222681
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 22 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 0 a 4 c 4 g 2 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 29
AR071794
LOCUS AR071794 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 23 from patent US 5912147.
ACCESSION AR071794
VERSION AR071794.1 GI:7222682

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 23 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 0 a 5 c 4 g 1 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 30
AR071794/c
LOCUS AR071794 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 23 from patent US 5912147.
ACCESSION AR071794
VERSION AR071794.1 GI:7222682
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 23 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 0 a 5 c 4 g 1 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 31
AR071795
LOCUS AR071795 10 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 24 from patent US 5912147.
ACCESSION AR071795
VERSION AR071795.1 GI:7222683
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability
JOURNAL Patent: US 5912147-A 24 15-JUN-1999;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"

BASE COUNT 0 a 4 c 4 g 2 t
ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;

Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 32
AR071795/c 10 bp DNA linear PAT 18-FEB-2000

LOCUS AR071795
DEFINITION Sequence 24 from patent US 5912147.

ACCESSION AR071795
VERSION AR071795.1 GI:7222683

KEYWORDS
SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability

JOURNAL Patent: US 5912147-A 24 15-JUN-1999;
FEATURES Location/Qualifiers

source 1..10
BASE COUNT 0 a 4 c 4 g 2 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 33

LOCUS AR071796 10 bp DNA linear PAT 18-FEB-2000

DEFINITION Sequence 25 from patent US 5912147.

ACCESSION AR071796
VERSION AR071796.1 GI:7222684

KEYWORDS
SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability

JOURNAL Patent: US 5912147-A 25 15-JUN-1999;
FEATURES Location/Qualifiers

source 1..10
BASE COUNT 0 a 5 c 4 g 1 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 34
AR071796/c 10 bp DNA linear PAT 18-FEB-2000

LOCUS AR071796
DEFINITION Sequence 25 from patent US 5912147.

ACCESSION AR071796
VERSION AR071796.1 GI:7222684

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 10)

AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability

JOURNAL Patent: US 5912147-A 25 15-JUN-1999;
FEATURES Location/Qualifiers

source 1..10
BASE COUNT 0 a 5 c 4 g 1 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 35

LOCUS AR071797 10 bp DNA linear PAT 18-FEB-2000

DEFINITION Sequence 26 from patent US 5912147.

ACCESSION AR071797
VERSION AR071797.1 GI:7222685

KEYWORDS
SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability

JOURNAL Patent: US 5912147-A 26 15-JUN-1999;
FEATURES Location/Qualifiers

source 1..10
BASE COUNT 0 a 4 c 4 g 2 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 36
AR071797/c 10 bp DNA linear PAT 18-FEB-2000

LOCUS AR071797
DEFINITION Sequence 26 from patent US 5912147.

ACCESSION AR071797
VERSION AR071797.1 GI:7222685

KEYWORDS
SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 10)

AUTHORS Stoler,D., Basik,M. and Anderson,G.
TITLE Rapid means of quantitating genomic instability

JOURNAL Patent: US 5912147-A 26 15-JUN-1999;
FEATURES Location/Qualifiers

source 1..10
BASE COUNT 0 a 4 c 4 g 2 t

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.9e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
 |||||||
 Db 8 CGCGCGCG 1

RESULT 37
 AR071798
 LOCUS AR071798 10 bp DNA linear PAT 18-FEB-2000
 DEFINITION Sequence 27 from patent US 5912147.
 ACCESSION AR071798
 VERSION AR071798.1 GI:7222686
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 10)
 AUTHORS Stoler,D., Basik,M. and Anderson,G.
 TITLE Rapid means of quantitating genomic instability
 JOURNAL Patent: US 5912147-A 27 15-JUN-1999;
 FEATURES Location/Qualifiers
 source 1..10
 BASE COUNT 0 a 4 c 4 g 2 t

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
 |||||||
 Db 1 CGCGCGCG 8

RESULT 38
 AR071798/c
 LOCUS AR071798 10 bp DNA linear PAT 18-FEB-2000
 DEFINITION Sequence 27 from patent US 5912147.
 ACCESSION AR071798
 VERSION AR071798.1 GI:7222686
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 10)
 AUTHORS Stoler,D., Basik,M. and Anderson,G.
 TITLE Rapid means of quantitating genomic instability
 JOURNAL Patent: US 5912147-A 27 15-JUN-1999;
 FEATURES Location/Qualifiers
 source 1..10
 BASE COUNT 0 a 4 c 4 g 2 t

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
 |||||||
 Db 8 CGCGCGCG 1

RESULT 39
 AR094565
 LOCUS AR094565 10 bp DNA linear PAT 08-SEP-2000
 DEFINITION Sequence 13 from patent US 6001657.
 ACCESSION AR094565
 VERSION AR094565.1 GI:10021597
 KEYWORDS
 SOURCE Unknown.

ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 10)
 AUTHORS Hardin,C.C., Brown,B.A. II, Roberts,J.F. and Pelsue,S.C.
 TITLE Antibodies that selectively bind quadruplex nucleic acids
 JOURNAL Patent: US 6001657-A 13 14-DEC-1999;
 FEATURES Location/Qualifiers
 source 1..10
 BASE COUNT 0 a 5 c 5 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
 |||||||
 Db 1 CGCGCGCG 8

RESULT 40
 AR094565/c
 LOCUS AR094565 10 bp DNA linear PAT 08-SEP-2000
 DEFINITION Sequence 13 from patent US 6001657.
 ACCESSION AR094565
 VERSION AR094565.1 GI:10021597
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 10)
 AUTHORS Hardin,C.C., Brown,B.A. II, Roberts,J.F. and Pelsue,S.C.
 TITLE Antibodies that selectively bind quadruplex nucleic acids
 JOURNAL Patent: US 6001657-A 13 14-DEC-1999;
 FEATURES Location/Qualifiers
 source 1..10
 BASE COUNT 0 a 5 c 5 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 8.9e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCGCG 8
 |||||||
 Db 10 CGCGCGCG 3

RESULT 41
 AR199370
 LOCUS AR199370 12 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 79 from patent US 6355428.
 ACCESSION AR199370
 VERSION AR199370.1 GI:20249444
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 12)
 AUTHORS Schroth,G.P., Brulice,T. Wayne, and Suh,Y.J.
 TITLE Nucleic acid ligand interaction assays
 JOURNAL Patent: US 6355428-A 79 12-MAR-2002;
 FEATURES Location/Qualifiers
 source 1..12
 BASE COUNT 0 a 8 c 4 g 0 t

Query Match 100.0%; Score 8; DB 6; Length 12;
 Best Local Similarity 100.0%; Pred. No. 8.6e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 DB 3 CGCGCGCG 10

RESULT 42

AR199370/c

LOCUS AR199370 12 bp DNA PAT 20-APR-2002
 DEFINITION Sequence 79 from patent US 6355428.
 ACCESSION AR199370
 VERSION AR199370.1 GI:20249444

KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 12)
 AUTHORS Schroth,G.P., Brulice,T. Wayne. and Suh,Y.J.
 TITLE Nucleic acid ligand interaction assays
 JOURNAL Patent: US 6355428-A 79 12-MAR-2002;
 FEATURES Location/Qualifiers

source 1..12
 /organism="unknown"
 BASE COUNT 0 a 8 c 4 g 0 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 12;
 Best Local Similarity 100.0%; Pred. No. 8.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 DB 10 CGCGCGCG 3

RESULT 43

AR048207

LOCUS AR048207 16 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 3 from patent US 5821070.
 ACCESSION AR048207
 VERSION AR048207.1 GI:5970550

KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 16)
 AUTHORS Lee,W.-H. and Shan,B.
 TITLE Antibodies reactive with retinoblastoma binding proteins and
 methods of using same
 JOURNAL Patent: US 5821070-A 3 13-OCT-1998;
 FEATURES Location/Qualifiers

source 1..16
 /organism="unknown"
 BASE COUNT 3 a 5 c 5 g 3 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 DB 6 CGCGCGCG 13

RESULT 44

AR048207/c

LOCUS AR048207 16 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 3 from patent US 5821070.
 ACCESSION AR048207
 VERSION AR048207.1 GI:5970550

KEYWORDS
 SOURCE Unknown.

ORGANISM Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 16)
 AUTHORS Lee,W.-H. and Shan,B.
 TITLE Antibodies reactive with retinoblastoma binding proteins and
 methods of using same
 JOURNAL Patent: US 5821070-A 3 13-OCT-1998;
 FEATURES Location/Qualifiers

source 1..16
 /organism="unknown"
 BASE COUNT 3 a 5 c 5 g 3 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 DB 11 CGCGCGCG 4

RESULT 45

AR055680

LOCUS AR055680 16 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 3 from patent US 5837512.
 ACCESSION AR055680
 VERSION AR055680.1 GI:5981257

KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 16)
 AUTHORS Rabson,A.B., Lin,H.-C., Bodkin,M. and Strahl,R.
 TITLE HIV-1 vectors
 JOURNAL Patent: US 5837512-A 3 17-NOV-1998;
 FEATURES Location/Qualifiers

source 1..16
 /organism="unknown"
 BASE COUNT 0 a 8 c 8 g 0 t
 ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 DB 8 CGCGCGCG 15

Search completed: March 14, 2003, 04:08:03
 Job time : 1364 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 01:45:32 ; Search time 197 seconds
(without alignments)
91.452 Million cell updates/sec

Title: CGCGCGCG

Perfect score: 8

Sequence: 1 cgcgcgcg 8

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	100.0	8	22	AA511907
2	8	100.0	8	22	AA511907
3	8	100.0	10	20	AA511907
4	8	100.0	10	20	AA511907
5	8	100.0	10	20	AA511907
6	8	100.0	10	20	AA511907
7	8	100.0	10	20	AA511907
8	100.0	10	20	AA511907	
9	8	100.0	10	20	AA511907

C	10	8	100.0	10	20	AA511907	US5912147 primer 2
C	11	8	100.0	10	20	AA511907	US5912147 primer 2
C	12	8	100.0	10	20	AA511907	US5912147 primer 2
C	13	8	100.0	10	20	AA511907	US5912147 primer 2
C	14	8	100.0	10	20	AA511907	US5912147 primer 2
C	15	8	100.0	10	20	AA511907	US5912147 primer 2
C	16	8	100.0	10	20	AA511907	US5912147 primer 2
C	17	8	100.0	10	20	AA511907	US5912147 primer 2
C	18	8	100.0	10	20	AA511907	US5912147 primer 2
C	19	8	100.0	10	20	AA511907	US5912147 primer 2
C	20	8	100.0	10	20	AA511907	US5912147 primer 2
C	21	8	100.0	10	20	AA511907	US5912147 primer 2
C	22	8	100.0	10	20	AA511907	US5912147 primer 2
C	23	8	100.0	10	20	AA511907	US5912147 primer 2
C	24	8	100.0	10	20	AA511907	US5912147 primer 2
C	25	8	100.0	10	20	AA511907	US5912147 primer 2
C	26	8	100.0	10	20	AA511907	US5912147 primer 2
C	27	8	100.0	10	20	AA511907	US5912147 primer 2
C	28	8	100.0	10	20	AA511907	US5912147 primer 2
C	29	8	100.0	10	20	AA511907	US5912147 primer 2
C	30	8	100.0	10	20	AA511907	US5912147 primer 2
C	31	8	100.0	10	20	AA511907	US5912147 primer 2
C	32	8	100.0	10	20	AA511907	US5912147 primer 2
C	33	8	100.0	10	20	AA511907	US5912147 primer 2
C	34	8	100.0	10	20	AA511907	US5912147 primer 2
C	35	8	100.0	10	20	AA511907	US5912147 primer 2
C	36	8	100.0	10	20	AA511907	US5912147 primer 2
C	37	8	100.0	10	20	AA511907	US5912147 primer 2
C	38	8	100.0	10	20	AA511907	US5912147 primer 2
C	39	8	100.0	10	20	AA511907	US5912147 primer 2
C	40	8	100.0	10	20	AA511907	US5912147 primer 2
C	41	8	100.0	10	20	AA511907	US5912147 primer 2
C	42	8	100.0	10	20	AA511907	US5912147 primer 2
C	43	8	100.0	10	20	AA511907	US5912147 primer 2
C	44	8	100.0	10	20	AA511907	US5912147 primer 2
C	45	8	100.0	10	20	AA511907	US5912147 primer 2

ALIGNMENTS

RESULT 1
ID: AA511907 standard; DNA: 8 BP.

AC: AA511907;

DT: 07-NOV-2001 (first entry)

DE: Hypothetical forward PCR primer.

KW: Hypothetical PCR primer: DNA sequencing; cDNA cloning; single-stranded DNA amplification; circular template; ss.

OS: Synthetic.

PN: WO200159101-A1.

PD: 16-AUG-2001.

PF: 09-FEB-2001; 2001WO-US04259.

PR: 10-FEB-2000; 2000US-0181615.

PR: 09-MAY-2000; 2000US-0203035.

(PENN-) PENN STATE RES FOUND.

Connor JR, Ye Z.

WPI: 2001-522475/57.

Amplifying a polynucleotide sequence for molecular biology protocols, comprises circularizing a linear, single strand polynucleotide sample

PT and initiating primer extension amplification using sequence specific
 PT primers -
 XX Disclosure; Page 8; 39pp; English.
 PS
 CC The invention relates to methods of amplifying a polynucleotide sequence,
 CC involving obtaining a linear, single strand polynucleotide sample,
 CC ligating the ends of the sample to form a circular shaped sample,
 CC introducing first and second sequence specific primers to the circular
 CC sample and initiating a primer extension amplification reaction to
 CC increase copy number of the circular sample. The new methods are useful
 CC for amplifying a polynucleotide sequence, cloning a full length cDNA
 CC sequence from an mRNA sample and sequencing a full length coding DNA or
 CC mRNA for a gene (cloned). The products obtained are useful for a number
 CC of molecular biology protocols including diagnostics, sequencing and
 CC mutation. The method uses a single-strand of polynucleotides and can
 CC amplify the single strand cDNA obtained after reverse transcription of
 CC mRNA rather than double-stranded cDNA, hence the method is accurate and
 CC efficient in amplification. The method is less time consuming and
 CC inexpensive. Background of polymerase chain reaction (PCR) products is
 CC reduced as specific primers are used in all amplification reactions and
 CC as the primer binds to both ends of the cDNA, and even low expressed
 CC mRNA can be cloned. Since the method uses only first strand cDNA as a
 CC PCR template, the longest first strand cDNA could be synthesised by
 CC using reverse transcriptase without RNase H activity. The cDNA band of
 CC correct size can be obtained on the first pass or the amplification of
 CC cDNA ends can be repeated until the correct size of the cDNA is
 CC obtained. The amplification of cDNA towards the ends, which is contrary
 CC to normal gene structure, decreases the possibility of contamination in
 CC cDNA cloning from genomic DNA. The technique can also be used as a
 CC better alternative to existing methods for 5' end primer extension
 CC because of its ability to specifically amplify cDNA ends using a grade
 CC series of amplification steps. The present sequence is a hypothetical
 CC PCR primer used to illustrate the method of the invention.
 CC
 SQ Sequence 8 BP; 0 A; 4 C; 4 G; 0 U; 0 other;
 Query Match 100.0%; Score 8; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+08;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCGCGCG 8
 Db 1 CGCGCGCG 8
 RESULT 2
 AAS11907/C
 ID AAS11907 standard; DNA; 8 BP.
 XX
 AC AAS11907;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Hypothetical forward PCR primer.
 XX
 KW Hypothetical PCR primer; DNA sequencing; cDNA cloning;
 KW single-stranded DNA amplification; circular template; ss.
 XX
 OS Synthetic.
 XX
 PN WO200159101-A1.
 PD 16-AUG-2001.
 XX
 PF 09-FEB-2001; 2001WO-US04259.
 XX
 PR 10-FEB-2000; 2000US-0181615.
 XX
 PR 09-MAY-2000; 2000US-0203035.
 XX
 PA (PENN-) PENN STATE RES FOUND.
 XX
 PI Connor JR, Ye Z;

XX
 DR WPI; 2001-522475/57.
 XX
 PT Amplifying a polynucleotide sequence for molecular biology protocols,
 PT comprises circularizing a linear, single strand polynucleotide sample
 PT and initiating primer extension amplification using sequence specific
 PT primers -
 XX Disclosure; Page 8; 39pp; English.
 PS
 CC The invention relates to methods of amplifying a polynucleotide sequence,
 CC involving obtaining a linear, single strand polynucleotide sample,
 CC ligating the ends of the sample to form a circular shaped sample,
 CC introducing first and second sequence specific primers to the circular
 CC sample and initiating a primer extension amplification reaction to
 CC increase copy number of the circular sample. The new methods are useful
 CC for amplifying a polynucleotide sequence, cloning a full length cDNA
 CC sequence from an mRNA sample and sequencing a full length coding DNA or
 CC mRNA for a gene (cloned). The products obtained are useful for a number
 CC of molecular biology protocols including diagnostics, sequencing and
 CC mutation. The method uses a single-strand of polynucleotides and can
 CC amplify the single strand cDNA obtained after reverse transcription of
 CC mRNA rather than double-stranded cDNA, hence the method is accurate and
 CC efficient in amplification. The method is less time consuming and
 CC inexpensive. Background of polymerase chain reaction (PCR) products is
 CC reduced as specific primers are used in all amplification reactions and
 CC as the primer binds to both ends of the cDNA, and even low expressed
 CC mRNA can be cloned. Since the method uses only first strand cDNA as a
 CC PCR template, the longest first strand cDNA could be synthesised by
 CC using reverse transcriptase without RNase H activity. The cDNA band of
 CC correct size can be obtained on the first pass or the amplification of
 CC cDNA ends can be repeated until the correct size of the cDNA is
 CC obtained. The amplification of cDNA towards the ends, which is contrary
 CC to normal gene structure, decreases the possibility of contamination in
 CC cDNA cloning from genomic DNA. The technique can also be used as a
 CC better alternative to existing methods for 5' end primer extension
 CC because of its ability to specifically amplify cDNA ends using a grade
 CC series of amplification steps. The present sequence is a hypothetical
 CC PCR primer used to illustrate the method of the invention.
 CC
 SQ Sequence 8 BP; 0 A; 4 C; 4 G; 0 U; 0 other;
 Query Match 100.0%; Score 8; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+08;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCGCGCG 8
 Db 8 CGCGCGCG 1
 RESULT 3
 AAX77475
 ID AAX77475 standard; DNA; 10 BP.
 XX
 AC AAX77475;
 XX
 DT 05-AUG-1999 (first entry)
 XX
 DE US5912147 primer 19.
 XX
 KW Primer; quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; ss.
 XX
 OS Synthetic.
 XX
 PN US5912147-A.
 PD 15-JUN-1999.
 XX
 PF 22-OCT-1996; 96US-0734973.
 XX
 PR 22-OCT-1996; 96US-0734973.
 XX
 PR 22-OCT-1996; 96US-0734973.

XX
PA (HEAL-) HEALTH RES INC.
BY

PI Anderson G, Basik M, Stoler D;
 XY

DR WPI; 1999-357197/30.
XX

Pf Quantitating genetic instability
XX

PS CLAIM 4; Column 23-24; 27pp; English.
XX

has invention described a novel method for quantitating genetic instability independent of microsatellite alterations by treating a comparison pair comprising genomic DNA from tumour cells and genomic DNA from normal cells. The method involves the cells from the same individual with oligonucleotide primers selected from (i) a nucleotide sequence (CA)XRg, where R is a purine selected from adenine and guanine and x - 3-7, (ii) a nucleotide sequence (Co)xyR, where R is as in (i) and y is a pyrimidine selected from cytosine, thymine, and uracil and x - 3-7, (iii) a nucleotide sequence (Co)xyR, where R is as in (i) and x - 3-7, (iv) a nucleotide sequence (Co)xyY, where y is a pyrimidine selected from cytosine, thymine, and uracil and x - 3-7, (v) a nucleotide sequence (CA)XRg, where R is a purine selected from adenine and guanine and x - 6-16, (vi) a nucleotide sequence (CA)xyR, where R is a purine selected from adenine and guanine and y is a pyrimidine selected from cytosine, thymine, and uracil, and x - 6-16, (vii) a nucleotide sequence (CA)XRr, where R is a purine selected from adenine and guanine and x - 6-16, (viii) a nucleotide sequence (CA)xyY, where y is a pyrimidine selected from cytosine, thymine, and uracil and x - 6-16, and (ix) a combination of the primers. The method is useful for detecting genomic instability which are commonly associated with the various stages of neoplastic transformation and carcinogenesis. The method is rapid and simple.

Sequence 10 BP; 0 A; 6 C; 4 G; 0 U; 0 other;

Query Match	100.0%	Score 8;	DB 20;	Length 10;
Best Local Similarity	100.0%;	Pred. No. 5.3e+04;		
Matches	8;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0

OY	1	CGCGCGCG	8
Db	1	CGCGCGCG	8

RESULT 4
AAX77475/C

ID AAX77475 standard; DNA; 10 BP.

AC AAX77475;

DT 05-AUG-1999 (first entry)
YY

DE SS5912147 primer 19.
yy

primers; quantitation; genetic instability; tumour cell; detection; neoplastic transformation; condensation

XX
OS Synthetic

AA
PN US5912147-A.

PD 15-JUN-1999

PF 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

PA (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;
XY

WPI; 1999-357197/30.

PT Quantitating genetic instability
XX
PS Claim 4; Column 23-24; 27pp; English
XX

This invention describes a novel method for quantitating genetic instability independent of microsatellite alterations by treating a comparison pair comprising genomic DNA from tumour cells and genomic DNA from normal cells. The method involves the cells from the same individual with oligonucleotide primers selected from (i) a nucleotide sequence (CG)_xXG, where R is a purine selected from adenine and guanine and x = 3-7, (ii) a nucleotide sequence (CC)_yXY, where R is as in (i) and Y is a pyrimidine selected from cytosine, thymine, and uracil and X is a nucleotide sequence (CC)_xXR, where R is as in (i) and x = 3-7, (iii) a nucleotide sequence (CG)_xYV, where Y is as in (i) and x = 3-7, (iv) a cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence (CA)_xXG, where R is a purine selected from adenine and guanine and x = 6-16, (vi) a nucleotide sequence (CA)_xRY, where R is a purine selected from adenine and guanine and Y is a pyrimidine selected from cytosine, thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)_xR, where R is a purine selected from adenine and guanine and x = 6-16, (viii) a nucleotide sequence (CA)_xYV, where Y is as in (i) and x = 6-16, (ix) a combination from cytosine, thymine, and uracil and x = 6-16, and (x) a combination of the primers. The method is useful for detecting genomic instability which are commonly associated with the various stages of neoplastic transformation and carcinogenesis. The method is rapid and simple.

sq Sequence 10 BP; 0 A; 6 C; 4 G; 0 U; 0 other;

Query Match	100.0%;	Score 8;	DB 20;	Length 10;
Best Local Similarity	100.0%;	Pred. NO. 5.3e+04;		
Matches	8;	Mismatches	0;	Indels 0;
				Gaps 0;

QY	1	CGCGGCGG	8
Db	8	CGCGGCGG	1

RESULT 5
AAX77476

ID	AAX77476 standard; DNA; 10 BP.
XY	

AC AAX77476;
XX

05-AUG-1999 (first entry)

DE US391214 / primer 20.
XX

KW neoplastic transformation: carcinogenesis: ss
KW PI primer; quantification; genetic instability; tumour cell; detection;

OS Synthetic.

PN US5912147-A.

PD 15-JUN-1999

PF 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.
YY

PA (HEAL-) HEALTH RES INC.
XX

XX Anderson G, Basik M, Stoler D,

WFL; 1999-33/19/30
XX

XX
XX
generally
insubstantially

PS Claim 4; Column 23-24; 27pp; English.

CC This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA

CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX
 SQ Sequence 10 BP; 0 A; 5 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 11111111
 DB 1 CGCGCGCG 8

RESULT 6

AA77476/c
 ID AAX77476 standard; DNA; 10 BP.

XX AAX77476;

DT 05-AUG-1999 (first entry)

DE US5912147 primer 20.

XX Primer: quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; ss.

OS Synthetic.

XX US5912147-A.

PN 15-JUN-1999.

XX 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

PA (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;

DR WPI; 1999-357197/30.

XX Quantitating genetic instability

PS Claim 4; Column 23-24; 27pp; English.

CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from

CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX
 SQ Sequence 10 BP; 0 A; 5 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 11111111
 DB 8 CGCGCGCG 1

RESULT 7
 AAX77477
 ID AAX77477 standard; DNA; 10 BP.

XX AAX77477;

DT 05-AUG-1999 (first entry)

DE US5912147 primer 21.

XX Primer: quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.

OS Synthetic.

XX Key

FT misc_RNA

FT 10

FT /tag= a

FT /note= "uracil"

PN US5912147-A.

PD 15-JUN-1999.

XX 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

PA (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;

DR WPI; 1999-357197/30.

XX Quantitating genetic instability

PS Claim 4; Column 25-26; 27pp; English.

CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =

XX Sequence 10 BP; 0 A; 4 C; 4 G; 2 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04; Mismatches 0; Indels 0; Gaps 0;

DB 1 CGCGCGCG 8
1 CGCGCGCG 8

RESULT 10
AAX77478/c
ID AAX77478 standard; DNA; 10 BP.

AC AAX77478;

XX 05-AUG-1999 (first entry)

DE US5912147 primer 22.

KM Primer; quantitation; genetic instability; tumour cell; detection;
neoplastic transformation; carcinogenesis; ss.

XX Synthetic.

PN US5912147-A.

PD 15-JUN-1999.

PE 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

PA (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;

DR WPI; 1999-357197/30.

PT Quantitating genetic instability

PS Claim 4; Column 25-26; 27pp; English.

XX This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and Y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 0 A; 4 C; 4 G; 2 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
DB 8 CGCGCGCG 1

RESULT 11
AAX77479
ID AAX77479 standard; DNA; 10 BP.

AC AAX77479;

XX 05-AUG-1999 (first entry)

DE US5912147 primer 23.

KM Primer; quantitation; genetic instability; tumour cell; detection;
neoplastic transformation; carcinogenesis; ss.

XX Synthetic.

PN US5912147-A.

PD 15-JUN-1999.

PE 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

PA (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;

DR WPI; 1999-357197/30.

PT Quantitating genetic instability

PS Claim 4; Column 25-26; 27pp; English.

XX This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and Y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 0 A; 5 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
DB 1 CGCGCGCG 8

RESULT 12
AAX77479/c

```

ID AAX77479 standard; DNA: 10 BP.
XX
AC AAX77479;
XX
DT 05-AUG-1999 (first entry)
XX
DE US5912147 primer 23.
XX
KW Primer: quantitation; genetic instability; tumour cell; detection;
KW neoplastic transformation; carcinogenesis; ss.
XX
OS Synthetic.
XX
US5912147-A.
XX
PN 15-JUN-1999.
XX
PD 15-JUN-1999.
XX
PF 22-OCT-1996; 96US-0734973.
XX
PR 22-OCT-1996; 96US-0734973.
XX
PA (HEAL-) HEALTH RES INC.
XX
PI Anderson G, Basik M, Stoler D;
XX
DR WPI: 1999-357197/30.
XX
PT Quantitating genetic instability
XX
PS Claim 4; Column 25-26; 27pp; English.
XX
CC This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)XY, where R is as in (i) and Y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vi) a nucleotide sequence (CA)XY, where R is a purine selected
CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.
XX
SO Sequence 10 BP; 0 A; 5 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 13
AAX77481
ID AAX77481 standard; DNA: 10 BP.
XX
AC AAX77481;
XX
DT 05-AUG-1999 (first entry)
XX
DE US5912147 primer 25.

```

```

XX
KW Primer: quantitation; genetic instability; tumour cell; detection;
KW neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT msc_RNA 9
FT /*tag= a
FT /note="uracil"

US5912147-A.
XX
PN 15-JUN-1999.
XX
PD 15-JUN-1999.
XX
PF 22-OCT-1996; 96US-0734973.
XX
PR 22-OCT-1996; 96US-0734973.
XX
PA (HEAL-) HEALTH RES INC.
XX
PI Anderson G, Basik M, Stoler D;
XX
DR WPI: 1999-357197/30.
XX
PT Quantitating genetic instability
XX
PS Claim 4; Column 25-26; 27pp; English.
XX
CC This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)XY, where R is as in (i) and Y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vi) a nucleotide sequence (CA)XY, where R is a purine selected
CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.
XX
SO Sequence 10 BP; 0 A; 5 C; 4 G; 1 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 14
AAX77481/C
ID AAX77481 standard; DNA: 10 BP.
XX
AC AAX77481;
XX
DT 05-AUG-1999 (first entry)
XX
DE US5912147 primer 25.
XX
KW Primer: quantitation; genetic instability; tumour cell; detection;

```

```

KM      neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
XX
XX      Synthetic.
OS
FH      Key
FH      Location/Qualifiers
FT      misc-RNA
FT      9
FT      /*tag= a
FT      /note= "uracil"
XX
XX      US5912147-A.
XX
XX      15-JUN-1999.
XX
XX      22-OCT-1996; 96US-0734973.
XX
XX      22-OCT-1996; 96US-0734973.
XX
XX      (HEAL-) HEALTH RES INC.
XX
XX      Anderson G, Basik M, Stoler D;
XX      WPI; 1999-357197/30.
XX
XX      Quantitating genetic instability
XX
XX      Claim 4; Column 25-26; 27pp; English.
XX
XX      This invention describes a novel method for quantitating genetic
XX      instability independent of microsatellite alterations by treating a
XX      comparison pair comprising genomic DNA from tumour cells and genomic DNA
XX      from normal cells. The method involves the cells from the same individual
XX      with oligonucleotide primers selected from (i) a nucleotide sequence
XX      (CG)XRG, where R is a purine selected from adenine and guanine and x =
XX      3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and y is a
XX      pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
XX      a nucleotide sequence (CG)XYR, where R is as in (i) and x = 3-7, (iv) a
XX      nucleotide sequence (CG)XY, where y is a pyrimidine selected from
XX      cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
XX      (CA)XRG, where R is a purine selected from adenine and guanine and x =
XX      6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
XX      from adenine and guanine and y is a pyrimidine selected from cytosine,
XX      thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
XX      where R is a purine selected from adenine and guanine and x = 6-16,
XX      (viii) a nucleotide sequence (CA)XYR, where y is a pyrimidine selected
XX      from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
XX      of the primers. The method is useful for detecting genomic instability
XX      which are commonly associated with the various stages of neoplastic
XX      transformation and carcinogenesis. The method is rapid and simple.
XX
XX      Sequence 10 BP; 0 A; 5 C; 4 G; 1 U; 0 other;
XX
XX      Query Match
XX      Best Local Similarity 100.0%; Score 8; DB 20; Length 10;
XX      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY      1 CGCGCGCG 8
XX      |||||||
XX      8 CGCGCGCG 1
XX
XX      RESULT 15
XX      ID AAX77480
XX      AAX77480 standard; DNA; 10 BP.
XX
XX      AAX77480;
XX
XX      05-AUG-1999 (first entry)
XX
XX      US5912147 primer 24.
XX
XX      Primer: quantitation; genetic instability; tumour cell; detection;
XX      neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
XX

```

```

OS      Synthetic.
XX
XX      Key
XX      Location/Qualifiers
FH      misc-RNA
FH      10
FH      /*tag= a
FH      /note= "uracil"
XX
XX      US5912147-A.
XX
XX      15-JUN-1999.
XX
XX      22-OCT-1996; 96US-0734973.
XX
XX      22-OCT-1996; 96US-0734973.
XX
XX      (HEAL-) HEALTH RES INC.
XX
XX      Anderson G, Basik M, Stoler D;
XX      WPI; 1999-357197/30.
XX
XX      Quantitating genetic instability
XX
XX      Claim 4; Column 25-26; 27pp; English.
XX
XX      This invention describes a novel method for quantitating genetic
XX      instability independent of microsatellite alterations by treating a
XX      comparison pair comprising genomic DNA from tumour cells and genomic DNA
XX      from normal cells. The method involves the cells from the same individual
XX      with oligonucleotide primers selected from (i) a nucleotide sequence
XX      (CG)XRG, where R is a purine selected from adenine and guanine and x =
XX      3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and y is a
XX      pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
XX      a nucleotide sequence (CG)XYR, where R is as in (i) and x = 3-7, (iv) a
XX      nucleotide sequence (CG)XY, where y is a pyrimidine selected from
XX      cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
XX      (CA)XRG, where R is a purine selected from adenine and guanine and x =
XX      6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
XX      from adenine and guanine and y is a pyrimidine selected from cytosine,
XX      thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
XX      where R is a purine selected from adenine and guanine and x = 6-16,
XX      (viii) a nucleotide sequence (CA)XYR, where y is a pyrimidine selected
XX      from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
XX      of the primers. The method is useful for detecting genomic instability
XX      which are commonly associated with the various stages of neoplastic
XX      transformation and carcinogenesis. The method is rapid and simple.
XX
XX      Sequence 10 BP; 0 A; 4 C; 4 G; 1 T; 1 U; 0 other;
XX
XX      Query Match
XX      Best Local Similarity 100.0%; Score 8; DB 20; Length 10;
XX      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY      1 CGCGCGCG 8
XX      |||||||
XX      1 CGCGCGCG 8
XX
XX      RESULT 16
XX      ID AAX77480/c
XX      AAX77480;
XX
XX      05-AUG-1999 (first entry)
XX
XX      US5912147 primer 24.
XX
XX      Primer: quantitation; genetic instability; tumour cell; detection;
XX      neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
XX
XX      Synthetic.
XX

```

FT	Key	Location/Qualifiers
FT	misc_RNA	10
FT		/tag= a
FT		/note= "uracil"
PN	US5912147-A.	
XX		
PD	15-JUN-1999.	
XX		
PF	22-OCT-1996;	96US-0734973.
XX		
PR	22-OCT-1996;	96US-0734973.
XX		
PA	(HEAL-) HEALTH RES INC.	
XX		
PI	Anderson G, Basik M, Stoler D,	
XX		
DR	WPI; 1999-357197/30.	
XX		
PT		
XX		
PS	Quantitating genetic instability	
XX		
CC	Claim 4; Column 25-26; 27pp; English.	
XX		
CC	This invention describes a novel method for quantitating genetic	
CC	instability independent of microsatellite alterations by treating a	
CC	comparison pair comprising genomic DNA from tumour cells and genomic DNA	
CC	from normal cells. The method involves the cells from the same individual	
CC	with oligonucleotide primers selected from (1) a nucleotide sequence	
CC	(CG)xRG, where R is a purine selected from adenine and guanine and x =	
CC	3-7, (11) a nucleotide sequence (CG)xRY, where R is as in (1) and Y is a	
CC	pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (111)	
CC	a nucleotide sequence (CG)xRR, where R is as in (1) and x = 3-7, (1v) a	
CC	nucleotide sequence (CG)xYX, where Y is a pyrimidine selected from	
CC	cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence	
CC	(CA)xRG, where R is a purine selected from adenine and guanine and x =	
CC	6-16, (vi) a nucleotide sequence (CA)xRY, where R is a purine selected	
CC	from adenine and guanine and Y is a pyrimidine selected from cytosine,	
CC	thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xRR,	
CC	where R is a purine selected from adenine and guanine and x = 6-16,	
CC	(viii) a nucleotide sequence (CA)xYX, where Y is a pyrimidine selected	
CC	from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination	
CC	of the primers. The method is useful for detecting genomic instability	
CC	which are commonly associated with the various stages of neoplastic	
CC	transformation and carcinogenesis. The method is rapid and simple.	
XX		
SO	Sequence 10 BP; 0 A; 4 C; 4 G; 1 T; 1 U; 0 other;	
XX		
Query Match	100.0%; Score 8; DB 20; Length 10;	
Best Local Similarity	100.0%; Pred. No. 5.3e+04;	
Matches 8; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
0Y	1 CGCGCGCG 8	
Db	8 CGCGCGCG 1	
XX		
RESULT 17		
AAK77482		
ID	AAK77482 standard; DNA: 10 BP.	
XX		
AC	AAK77482;	
XX		
DT	05-AUG-1999 (first entry)	
XX		
DE	US5912147 primer 26.	
XX		
PR	Primer; quantitation; genetic instability; tumour cell; detection;	
KW	neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.	
XX		
OS	Synthetic.	
XX		
Key		
FT	9	Location/Qualifiers
misc_RNA		

FT	/*tag= a
FT	/note= "uracil"
XX	
XX	US5912147-A.
XX	
PN	15-JUN-1999.
PD	
XX	
PF	22-OCT-1996; 9605-0734973.
XX	
PR	22-OCT-1996; 9605-0734973.
XX	
PA	(HEAL-) HEALTH RES INC.
PI	Anderson G, Basik M, Stoler D;
XX	
DR	WPI; 1999-357197/30.
XX	
PT	Quantitating genetic instability
XX	
PS	Claim 4; Column 27-28; 27pp; English.
XX	
CC	This invention describes a novel method for quantitating genetic
CC	instability independent of microsatellite alterations by treating a
CC	comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC	from normal cells. The method involves the cells from the same individual
CC	with oligonucleotide primers selected from (i) a nucleotide sequence
CC	(CG)xRG, where R is a purine selected from adenine and guanine and x =
CC	3-7, (ii) a nucleotide sequence (CG)xRY, where R is as in (i) and y is a
CC	pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC	a nucleotide sequence (CG)xRR, where R is as in (i) and x = 3-7, (iv) a
CC	nucleotide sequence (CG)xY, where Y is a pyrimidine selected from
CC	cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC	(CA)xRG, where R is a purine selected from adenine and guanine and x =
CC	6-16, (vi) a nucleotide sequence (CA)xRY, where R is a purine selected
CC	from adenine and guanine and y is a pyrimidine selected from cytosine,
CC	thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xRR,
CC	where R is a purine selected from adenine and guanine and x = 6-16,
CC	(viii) a nucleotide sequence (CA)xY, where Y is a pyrimidine selected
CC	from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC	of the primers. The method is useful for detecting genomic instability
CC	which are commonly associated with the various stages of neoplastic
CC	transformation and carcinogenesis. The method is rapid and simple.
XX	
SQ	Sequence 10 BP; 0 A; 4 C; 4 G; 1 T; 1 U; 0 other;
XX	
Query Match	100.0%; Score 8; DB 20; Length 10;
Best Local Similarity	100.0%; Pred. No. 5.3e+04;
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CGCGCGCG 8
DB	1 CGCGCGCG 8
XX	
RESULT 18	
ID	AAx77482/C
XX	
AC	AAx77482;
XX	
DT	05-AUG-1999 (first entry)
XX	
DE	US5912147 primer 26.
XX	
KW	Primer; quantitation; genetic instability; tumour cell; detection;
XX	neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
OS	Synthetic.
XX	
XX	
Key	Location/Qualifiers
FT	9
FT	/*tag= a
FT	/note= "uracil"

```

XX PN US5912147-A.
XX PD 15-JUN-1999.
XX PF 22-OCT-1996; 96US-0734973.
XX PR 22-OCT-1996; 96US-0734973.
XX PA (HEAL-) HEALTH RES INC.
XX PI Anderson G, Basik M, Stoler D:
XX DR WPI; 1999-357197/30.
XX PT Quantitating genetic instability
XX PS Claim 4; Column 27-28; 27pp; English.
XX CC This invention describes a novel method for quantitating genetic
XX CC instability independent of microsatellite alterations by treating a
XX CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
XX CC from normal cells. The method involves the cells from the same individual
XX CC with oligonucleotide primers selected from (i) a nucleotide sequence
XX CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
XX CC 3-7, (ii) a nucleotide sequence (CG)XR, where R is as in (i) and y is a
XX CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
XX CC a nucleotide sequence (CG)XY, where Y is as in (i) and x = 3-7, (iv) a
XX CC nucleotide sequence (CG)XR, where R is a pyrimidine selected from
XX CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
XX CC (CA)XR, where R is a purine selected from adenine and guanine and x =
XX CC 6-16, (vi) a nucleotide sequence (CA)XR, where R is a purine selected
XX CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
XX CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XR,
XX CC where R is a purine selected from adenine and guanine and x = 6-16,
XX CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
XX CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
XX CC of the primers. The method is useful for detecting genomic instability
XX CC which are commonly associated with the various stages of neoplastic
XX CC transformation and carcinogenesis. The method is rapid and simple.
XX SQ Sequence 10 BP; 0 A; 4 C; 4 G; 1 T; 1 U; 0 other:
XX
XX Query Match 100.0%; Score 8; DB 20; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 5.3e+04;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

```

```

RESULT 19
AA77483
ID AAX77483 standard; DNA; 10 BP.
XX AC AAX77483;
XX DT 05-AUG-1999 (first entry)
XX DE US5912147 primer 27.
XX KW Primer; quantitation; genetic instability; tumour cell; detection;
XX KW neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT 9..10
XX FT misc_RNA /*tag= a
XX FT /*note= "uracil"
XX PN US5912147-A.

```

```

XX PD 15-JUN-1999.
XX PF 22-OCT-1996; 96US-0734973.
XX PR 22-OCT-1996; 96US-0734973.
XX PA (HEAL-) HEALTH RES INC.
XX PI Anderson G, Basik M, Stoler D:
XX DR WPI; 1999-357197/30.
XX PT Quantitating genetic instability
XX PS Claim 4; Column 27-28; 27pp; English.
XX CC This invention describes a novel method for quantitating genetic
XX CC instability independent of microsatellite alterations by treating a
XX CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
XX CC from normal cells. The method involves the cells from the same individual
XX CC with oligonucleotide primers selected from (i) a nucleotide sequence
XX CC (CG)XR, where R is a purine selected from adenine and guanine and x =
XX CC 3-7, (ii) a nucleotide sequence (CG)XR, where R is as in (i) and y is a
XX CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
XX CC a nucleotide sequence (CG)XR, where R is a pyrimidine selected from
XX CC cytosine, thymine, and uracil and x = 3-7, (iv) a nucleotide sequence
XX CC (CA)XR, where R is a purine selected from adenine and guanine and x =
XX CC 6-16, (vi) a nucleotide sequence (CA)XR, where R is a purine selected
XX CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
XX CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XR,
XX CC where R is a purine selected from adenine and guanine and x = 6-16,
XX CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
XX CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
XX CC of the primers. The method is useful for detecting genomic instability
XX CC which are commonly associated with the various stages of neoplastic
XX CC transformation and carcinogenesis. The method is rapid and simple.
XX SQ Sequence 10 BP; 0 A; 4 C; 4 G; 2 U; 0 other:
XX
XX Query Match 100.0%; Score 8; DB 20; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 5.3e+04;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

```

```

RESULT 20
AA77483/c
ID AAX77483 standard; DNA; 10 BP.
XX AC AAX77483;
XX DT 05-AUG-1999 (first entry)
XX DE US5912147 primer 27.
XX KW Primer; quantitation; genetic instability; tumour cell; detection;
XX KW neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT 9..10
XX FT misc_RNA /*tag= a
XX FT /*note= "uracil"
XX PN US5912147-A.
XX PD 15-JUN-1999.

```

XX 22-OCT-1996; 96US-0734973.
 XX 22-OCT-1996; 96US-0734973.
 PR (HEAL-) HEALTH RES INC.
 XX
 XX
 PI Anderson G, Basik M, Stoler D;
 DR WPI; 1999-357197/30.
 XX
 PT Quantitating genetic instability
 PS Claim 4; Column 27-28; 27pp; English.
 XX
 CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XR_n, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XR_n, where R is as in (i) and y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XR_n, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XR_n, where R is as in (i) and x = 3-7, (iv) a
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XR_n, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XR_n, where R is a purine selected
 CC from adenine and guanine and y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XR_n,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XR_n, where y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.
 CC
 SQ Sequence 10 BP; 0 A; 4 C; 4 G; 2 U; 0 other;
 Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CGCGCGCG 8
 Db 8 CGCGCGCG 1
 RESULT 21
 AAX77465
 ID AAX77465 standard; DNA; 10 BP.
 XX
 AC AAX77465;
 XX
 DT 05-AUG-1999 (first entry)
 XX
 DE US5912147 primer 9.
 XX
 DE US5912147 primer 9.
 XX
 PN US5912147-A.
 XX
 PD 15-JUN-1999.
 XX
 PF 22-OCT-1996; 96US-0734973.
 XX
 PR 22-OCT-1996; 96US-0734973.
 XX
 PA (HEAL-) HEALTH RES INC.
 XX

PI Anderson G, Basik M, Stoler D;
 XX
 DR WPI; 1999-357197/30.
 XX
 PT Quantitating genetic instability
 PS Claim 4; Column 19-20; 27pp; English.
 XX
 CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XR_n, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XR_n, where R is as in (i) and y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XR_n, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XR_n, where R is as in (i) and x = 3-7, (iv) a
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XR_n, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XR_n, where R is a purine selected
 CC from adenine and guanine and y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XR_n,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XR_n, where y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.
 CC
 SQ Sequence 10 BP; 1 A; 4 C; 5 G; 0 U; 0 other;
 Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CGCGCGCG 8
 Db 1 CGCGCGCG 8
 RESULT 22
 AAX77465/c
 ID AAX77465 standard; DNA; 10 BP.
 XX
 AC AAX77465;
 XX
 DT 05-AUG-1999 (first entry)
 XX
 DE US5912147 primer 9.
 XX
 DE US5912147 primer 9.
 XX
 PN US5912147-A.
 XX
 PD 15-JUN-1999.
 XX
 PF 22-OCT-1996; 96US-0734973.
 XX
 PR 22-OCT-1996; 96US-0734973.
 XX
 PA (HEAL-) HEALTH RES INC.
 XX
 PI Anderson G, Basik M, Stoler D;
 XX
 DR WPI; 1999-357197/30.
 XX
 PT Quantitating genetic instability
 PS Claim 4; Column 19-20; 27pp; English.

XX This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRg, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XRy, where R is as in (i) and y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XRr, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XRy, where y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRg, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRy, where R is a purine selected
 CC from adenine and guanine and y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRr,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XRy, where R is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.
 CC

SQ Sequence 10 BP; 1 A; 4 C; 5 G; 0 U; 0 other;
 Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||
 Db 8 CGCGCGCG 1

RESULT 23
 AAX77466
 ID AAX77466 standard; DNA; 10 BP.
 XX AAX77466;
 AC
 XX 05-AUG-1999 (first entry)
 DT
 XX US5912147 primer 10.
 DE
 XX US5912147 primer 10.
 KW Primer: quantitation; genetic instability; tumour cell; detection;
 KM neoplastic transformation; carcinogenesis; ss.
 XX
 OS Synthetic.
 XX US5912147-A.
 PN
 XX 15-JUN-1999.
 PD
 XX 22-OCT-1996; 96US-0734973.
 PF
 XX 22-OCT-1996; 96US-0734973.
 PR
 XX (HEAL-) HEALTH RES INC.
 PA
 XX Anderson G, Basik M, Stoler D;
 PI
 XX WPI: 1999-357197/30.
 DR
 XX Quantitating genetic instability
 PT
 XX
 XX Claim 4; Column 19-20; 27pp; English.
 PS
 CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRg, where R is a purine selected from adenine and guanine and x =

CC 3-7, (ii) a nucleotide sequence (CG)XRy, where R is as in (i) and y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)XRr, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XRy, where y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRg, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRy, where R is a purine selected
 CC from adenine and guanine and y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRr,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XRy, where R is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.
 CC

SQ Sequence 10 BP; 0 A; 4 C; 6 G; 0 U; 0 other;
 Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||
 Db 1 CGCGCGCG 8

RESULT 24
 AAX77466/C
 ID AAX77466 standard; DNA; 10 BP.
 XX AAX77466;
 AC
 XX 05-AUG-1999 (first entry)
 DT
 XX US5912147 primer 10.
 DE
 XX US5912147 primer 10.
 KW Primer: quantitation; genetic instability; tumour cell; detection;
 KM neoplastic transformation; carcinogenesis; ss.
 XX
 OS Synthetic.
 XX US5912147-A.
 PN
 XX 15-JUN-1999.
 PD
 XX 22-OCT-1996; 96US-0734973.
 PF
 XX 22-OCT-1996; 96US-0734973.
 PR
 XX (HEAL-) HEALTH RES INC.
 PA
 XX Anderson G, Basik M, Stoler D;
 PI
 XX WPI: 1999-357197/30.
 DR
 XX Quantitating genetic instability
 PT
 XX
 XX Claim 4; Column 19-20; 27pp; English.
 PS
 CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRg, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XRy, where R is as in (i) and y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XRr, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XRy, where y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRg, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRy, where R is a purine selected

CC from adenine and guanine and y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)xYV, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.
 XX
 SO Sequence 10 BP; 0 A; 4 C; 6 G; 0 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 Db 8 CGCGCGCG 1

RESULT 25

AAx77467
 ID AAX77467 standard; DNA: 10 BP.

AC AAX77467;

DT 05-AUG-1999 (first entry)

XX US5912147 primer 11.

DE US5912147 primer 11.

XX Primer: quantitation; genetic instability; tumour cell; detection;

KW neoplastic transformation; carcinogenesis; ss.

XX Synthetic.

PN US5912147-A.

PD 15-JUN-1999.

XX 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

XX (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;

XX WPI; 1999-357197/30.

PT Quantitating genetic instability

XX Claim 4; Column 19-20; 27pp; English.

CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)xRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)xRY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)xYV, where Y is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)xYV, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)xRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)xRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)xYV, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic

CC transformation and carcinogenesis. The method is rapid and simple.
 XX
 SO Sequence 10 BP; 1 A; 5 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 Db 1 CGCGCGCG 8

RESULT 26

AAx77467/c
 ID AAX77467 standard; DNA: 10 BP.

AC AAX77467;

DT 05-AUG-1999 (first entry)

XX US5912147 primer 11.

DE US5912147 primer 11.

XX Primer: quantitation; genetic instability; tumour cell; detection;

KW neoplastic transformation; carcinogenesis; ss.

XX Synthetic.

PN US5912147-A.

PD 15-JUN-1999.

XX 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

XX (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;

XX WPI; 1999-357197/30.

PT Quantitating genetic instability

XX Claim 4; Column 19-20; 27pp; English.

CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)xRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)xRY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)xYV, where Y is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)xYV, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)xRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)xRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)xYV, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 |||||||
 Db 8 CGCGCGCG 1

RESULT 27
 AAX77468
 ID AAX77468 standard; DNA; 10 BP.
 XX
 AC AAX77468;
 XX
 DT 05-AUG-1999 (first entry)
 XX
 DE US5912147 primer 12.
 XX
 XX
 XX primer; quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; ss.
 XX
 OS Synthetic.
 XX
 PN US5912147-A.
 XX
 PD 15-JUN-1999.
 XX
 XX 22-OCT-1996; 96US-0734973.
 PF 22-OCT-1996; 96US-0734973.
 XX
 PR 22-OCT-1996; 96US-0734973.
 XX
 PA (HEAL-) HEALTH RES INC.
 XX
 PI Anderson G, Basik M, Stoler D;
 XX
 DR WPI; 1999-357197/30.
 XX

Quantitating genetic instability

Claim 4; Column 21-22; 27pp; English.

CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

SQ Sequence 10 BP; 1 A; 4 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 |||||||
 Db 1 CGCGCGCG 8

RESULT 28

AAX77468/c
 ID AAX77468 standard; DNA; 10 BP.
 XX
 AC AAX77468;
 XX
 DT 05-AUG-1999 (first entry)
 XX
 DE US5912147 primer 12.
 XX
 XX
 XX primer; quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; ss.
 XX
 OS Synthetic.
 XX
 PN US5912147-A.
 XX
 PD 15-JUN-1999.
 XX
 XX 22-OCT-1996; 96US-0734973.
 PF 22-OCT-1996; 96US-0734973.
 XX
 PR 22-OCT-1996; 96US-0734973.
 XX
 PA (HEAL-) HEALTH RES INC.
 XX
 PI Anderson G, Basik M, Stoler D;
 XX
 DR WPI; 1999-357197/30.
 XX

Quantitating genetic instability

Claim 4; Column 21-22; 27pp; English.

CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

SQ Sequence 10 BP; 1 A; 4 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 |||||||
 Db 8 CGCGCGCG 1

RESULT 29
 AAX77469
 ID AAX77469 standard; DNA; 10 BP.
 XX
 AC AAX77469;
 XX
 DT 05-AUG-1999 (first entry)
 XX

DE US5912147 primer 13.
 XX
 KW Primer: quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT misc_RNA 10
 FT /*tag= a
 FT /note= "uracil"
 FT
 PN US5912147-A.
 PD 15-JUN-1999.
 XX
 XX 22-OCT-1996; 96US-0734973.
 PF 22-OCT-1996; 96US-0734973.
 PR 22-OCT-1996; 96US-0734973.
 XX
 PA (HEAL-) HEALTH RES INC.
 XX
 PI Anderson G, Basik M, Stoler D;
 DR WPI: 1999-357197/30.
 XX
 PT Quantitating genetic instability

XX
 PS Claim 4; Column 21-22; 27pp; English.

XX This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (1) a nucleotide sequence
 CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (11) a nucleotide sequence (CG)XRY, where R is as in (1) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (111)
 CC a nucleotide sequence (CG)XRR, where R is as in (1) and x = 3-7, (111)
 CC cytosine, thymine, and uracil and x = 3-7, (1v) a
 CC nucleotide sequence (CG)XRY, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XRY, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 1 A; 4 C; 4 G; 1 U; 0 other;

QY Query Match 100.0%; Score 8; DB 20; Length 10;
 Db Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 CGCGCGCG 8
 XX |||||||
 Db 1 CGCGCGCG 8
 XX
 XX
 AC AAX77469;
 XX
 XX 05-AUG-1999 (first entry)
 DT
 XX
 DE US5912147 primer 13.
 XX

KW
 KW Primer: quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT misc_RNA 10
 FT /*tag= a
 FT /note= "uracil"
 FT
 PN US5912147-A.
 PD 15-JUN-1999.
 XX
 XX 22-OCT-1996; 96US-0734973.
 PF 22-OCT-1996; 96US-0734973.
 PR 22-OCT-1996; 96US-0734973.
 XX
 PA (HEAL-) HEALTH RES INC.
 XX
 PI Anderson G, Basik M, Stoler D;
 DR WPI: 1999-357197/30.
 XX
 PT Quantitating genetic instability

XX Claim 4; Column 21-22; 27pp; English.

XX This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (1) a nucleotide sequence
 CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (11) a nucleotide sequence (CG)XRY, where R is as in (1) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (111)
 CC a nucleotide sequence (CG)XRR, where R is as in (1) and x = 3-7, (111)
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XRY, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 1 A; 4 C; 4 G; 1 U; 0 other;

QY Query Match 100.0%; Score 8; DB 20; Length 10;
 Db Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 CGCGCGCG 8
 XX |||||||
 Db 8 CGCGCGCG 1
 XX
 XX
 AC AAX77470;
 XX
 XX 05-AUG-1999 (first entry)
 DT
 XX
 DE US5912147 primer 14.
 XX
 KW
 KW Primer: quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; ss.

XX Synthetic.
 OS US5912147-A.
 XX 15-JUN-1999.
 PD 22-OCT-1996; 96US-0734973.
 XX 22-OCT-1996; 96US-0734973.
 PR 22-OCT-1996; 96US-0734973.
 XX (HEAL-) HEALTH RES INC.
 PA Anderson G, Basik M, Stoler D;
 PI WPI; 1999-357197/30.
 DR Quantitating genetic instability
 PT Quantitating genetic instability
 PS Claim 4; Column 21-22; 27pp; English.

CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XRY, where y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
 CC from adenine and guanine and y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XRY, where y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 0 A; 5 C; 5 G; 0 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 DB 1 CGCGCGCG 8

RESULT 32

AAK77470/c

ID AAK77470 standard; DNA; 10 BP.

AC AAK77470;

DT 05-AUG-1999 (first entry)

DE US5912147 primer 14.

XX primer: quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; ss.

OS Synthetic.

PN US5912147-A.

PD 15-JUN-1999.

PF 22-OCT-1996; 96US-0734973.
 XX 22-OCT-1996; 96US-0734973.
 PR (HEAL-) HEALTH RES INC.
 PA Anderson G, Basik M, Stoler D;
 PI WPI; 1999-357197/30.
 DR Quantitating genetic instability
 PT Quantitating genetic instability
 PS Claim 4; Column 21-22; 27pp; English.

CC This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)XRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)XRY, where y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
 CC from adenine and guanine and y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)XRY, where y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 0 A; 5 C; 5 G; 0 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGCG 8
 DB 8 CGCGCGCG 1

RESULT 33

AAK77471

ID AAK77471 standard; DNA; 10 BP.

AC AAK77471;

DT 05-AUG-1999 (first entry)

DE US5912147 primer 15.

XX primer: quantitation; genetic instability; tumour cell; detection;
 KW neoplastic transformation; carcinogenesis; ss.

OS Synthetic.

PN US5912147-A.

PD 15-JUN-1999.

PF 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

PA (HEAL-) HEALTH RES INC.
 PI Anderson G, Basik M, Stoler D;

XX WPI: 1999-357197/30.

XX Quantitating genetic instability

PS Claim 4; Column 21-22; 27pp; English.

XX This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)xxR, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)xy, where R is as in (i) and y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)xxR, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)xy, where y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC (CA)xxR, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vi) a nucleotide sequence (CA)xy, where R is a purine selected
CC from adenine and guanine and y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xxR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)xy, where y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.

SO Sequence 10 BP; 0 A; 4 C; 5 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 34

AAx77471/c
ID AAX77471 standard; DNA: 10 BP.

XX AAX77471:

DT 05-AUG-1999 (first entry)

DE US5912147 primer 15.

KW Primer: quantitation; genetic instability; tumour cell; detection;
KM neoplastic transformation; carcinogenesis; ss.

OS Synthetic.

PN US5912147-A.

PD 15-JUN-1999.

PF 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

PA (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;

DR WPI: 1999-357197/30.

XX Quantitating genetic instability

PS Claim 4; Column 21-22; 27pp; English.

CC This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)xxR, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)xy, where R is as in (i) and y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)xxR, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)xy, where y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC (CA)xxR, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vi) a nucleotide sequence (CA)xy, where R is a purine selected
CC from adenine and guanine and y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xxR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)xy, where y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.

SO Sequence 10 BP; 0 A; 4 C; 5 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 35

AAx77472
ID AAX77472 standard; DNA: 10 BP.

XX AAX77472:

DT 05-AUG-1999 (first entry)

DE US5912147 primer 16.

KW Primer: quantitation; genetic instability; tumour cell; detection;
KM neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.

OS Synthetic.

FT Key Location/Qualifiers

FT misc_RNA 10
FT /tag= a
FT /note= "uracil"

PN US5912147-A.

PD 15-JUN-1999.

PF 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

PA (HEAL-) HEALTH RES INC.

PI Anderson G, Basik M, Stoler D;

DR WPI: 1999-357197/30.

XX Quantitating genetic instability

PS Claim 4; Column 23-24; 27pp; English.

CC This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a

CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)XY, where R is as in (i) and Y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)XRY, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vi) a nucleotide sequence (CA)XRY, where R is a purine selected
CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)XRR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.

XX
SQ Sequence 10 BP; 0 A; 4 C; 5 G; 1 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 36

AAX7472/c
ID AAX7472 standard; DNA: 10 BP.

XX AAX7472;

DT 05-AUG-1999 (first entry)

DE US5912147 primer 16.

KW Primer; quantitation; genetic instability; tumour cell; detection;
KW neoplastic transformation; carcinogenesis; DNA/RNA hybrid; ss.

XX OS Synthetic.

XX Key Location/Qualifiers

XX Key misc_RNA 10
XX FT /*tag= a
XX FT /note= "uracil"

XX US5912147-A.

PD 15-JUN-1999.

PF 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

XX (HEAL-) HEALTH RES INC.

XX Anderson G, Basik M, Stoler D;

XX WPI: 1999-357197/30.

XX Quantitating genetic instability

XX Claim 4; Column 23-24; 27pp; English.

CC This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual

CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)XRY, where R is as in (i) and Y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)XY, where Y is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)XRY, where R is as in (i) and x = 3-7, (v) a
CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (vi) a nucleotide sequence
CC (CA)XRG, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vii) a nucleotide sequence (CA)XRY, where R is a purine selected
CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (viii) a nucleotide sequence (CA)XRR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)XY, where Y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.

XX
SQ Sequence 10 BP; 0 A; 4 C; 5 G; 1 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 37

AAX7473
ID AAX7473 standard; DNA: 10 BP.

XX AAX7473;

DT 05-AUG-1999 (first entry)

DE US5912147 primer 17.

KW Primer; quantitation; genetic instability; tumour cell; detection;
KW neoplastic transformation; carcinogenesis; ss.

XX OS Synthetic.

XX US5912147-A.

PD 15-JUN-1999.

PF 22-OCT-1996; 96US-0734973.

PR 22-OCT-1996; 96US-0734973.

XX (HEAL-) HEALTH RES INC.

XX Anderson G, Basik M, Stoler D;

XX WPI: 1999-357197/30.

XX Quantitating genetic instability

XX Claim 4; Column 23-24; 27pp; English.

CC This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)XRG, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)XY, where R is as in (i) and Y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)XRY, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)XY, where Y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence

CC (CA)xRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)xRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)xYV, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 2 A; 4 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||
 Db 1 CGCGCGCG 8

RESULT 38
 AAX77473/c
 ID AAX77473 standard; DNA; 10 BP.

XX AAX77473;

XX 05-AUG-1999 (first entry)

XX US5912147 primer 17.

XX Primer: quantitation; genetic instability; tumour cell; detection;
 KM neoplastic transformation; carcinogenesis; ss.

XX Synthetic.

XX US5912147-A.

XX 15-JUN-1999.

XX 22-OCT-1996; 96US-0734973.

XX 22-OCT-1996; 96US-0734973.

XX (HEAL-) HEALTH RES INC.

XX Anderson G, Basik M, Stoler D;

XX WPI; 1999-357197/30.

XX Quantitating genetic instability

XX Claim 4; Column 23-24; 27pp; English.

XX This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)xRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)xRY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)xRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)xYV, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)xRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)xRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)xYV, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination

CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 2 A; 4 C; 4 G; 0 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
 |||||
 Db 8 CGCGCGCG 1

RESULT 39
 AAX77474
 ID AAX77474 standard; DNA; 10 BP.

XX AAX77474;

XX 05-AUG-1999 (first entry)

XX US5912147 primer 18.

XX Primer: quantitation; genetic instability; tumour cell; detection;
 KM neoplastic transformation; carcinogenesis; ss.

XX Synthetic.

XX US5912147-A.

XX 15-JUN-1999.

XX 22-OCT-1996; 96US-0734973.

XX 22-OCT-1996; 96US-0734973.

XX (HEAL-) HEALTH RES INC.

XX Anderson G, Basik M, Stoler D;

XX WPI; 1999-357197/30.

XX Quantitating genetic instability

XX Claim 4; Column 23-24; 27pp; English.

XX This invention describes a novel method for quantitating genetic
 CC instability independent of microsatellite alterations by treating a
 CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
 CC from normal cells. The method involves the cells from the same individual
 CC with oligonucleotide primers selected from (i) a nucleotide sequence
 CC (CG)xRG, where R is a purine selected from adenine and guanine and x =
 CC 3-7, (ii) a nucleotide sequence (CG)xRY, where R is as in (i) and Y is a
 CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
 CC a nucleotide sequence (CG)xRR, where R is as in (i) and x = 3-7, (iv) a
 CC nucleotide sequence (CG)xYV, where Y is a pyrimidine selected from
 CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
 CC (CA)xRG, where R is a purine selected from adenine and guanine and x =
 CC 6-16, (vi) a nucleotide sequence (CA)xRY, where R is a purine selected
 CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
 CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xRR,
 CC where R is a purine selected from adenine and guanine and x = 6-16,
 CC (viii) a nucleotide sequence (CA)xYV, where Y is a pyrimidine selected
 CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
 CC of the primers. The method is useful for detecting genomic instability
 CC which are commonly associated with the various stages of neoplastic
 CC transformation and carcinogenesis. The method is rapid and simple.

XX Sequence 10 BP; 1 A; 4 C; 5 G; 0 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 1 CGCGCGCG 8

RESULT 40

AAx77474/c
ID AAX77474 standard; DNA; 10 BP.

XX AAX77474;

XX 05-AUG-1999 (first entry)

DE US5912147 primer 18.

XX Primer; quantitation; genetic instability; tumour cell; detection;
KM neoplastic transformation; carcinogenesis; ss.

XX Synthetic.

XX US5912147-A.

XX 15-JUN-1999.

XX 22-OCT-1996; 96US-0734973.

XX 22-OCT-1996; 96US-0734973.

XX (HEAL-) HEALTH RES INC.

XX Anderson G, Basik M, Stoler D;

XX WPI; 1999-357197/30.

XX Quantitating genetic instability

PS Claim 4; Column 23-24; 27pp; English.

XX This invention describes a novel method for quantitating genetic
CC instability independent of microsatellite alterations by treating a
CC comparison pair comprising genomic DNA from tumour cells and genomic DNA
CC from normal cells. The method involves the cells from the same individual
CC with oligonucleotide primers selected from (i) a nucleotide sequence
CC (CG)xxR, where R is a purine selected from adenine and guanine and x =
CC 3-7, (ii) a nucleotide sequence (CG)xxY, where R is as in (i) and Y is a
CC pyrimidine selected from cytosine, thymine, and uracil and x = 3-7, (iii)
CC a nucleotide sequence (CG)xxR, where R is as in (i) and x = 3-7, (iv) a
CC nucleotide sequence (CG)xxY, where Y is a pyrimidine selected from
CC cytosine, thymine, and uracil and x = 3-7, (v) a nucleotide sequence
CC (CA)xxR, where R is a purine selected from adenine and guanine and x =
CC 6-16, (vi) a nucleotide sequence (CA)xxY, where R is a purine selected
CC from adenine and guanine and Y is a pyrimidine selected from cytosine,
CC thymine, and uracil, and x = 6-16, (vii) a nucleotide sequence (CA)xxR,
CC where R is a purine selected from adenine and guanine and x = 6-16,
CC (viii) a nucleotide sequence (CA)xxY, where Y is a pyrimidine selected
CC from cytosine, thymine, and uracil and x = 6-16, and (ix) a combination
CC of the primers. The method is useful for detecting genomic instability
CC which are commonly associated with the various stages of neoplastic
CC transformation and carcinogenesis. The method is rapid and simple.

SO Sequence 10 BP; 1 A; 4 C; 5 G; 0 U; 0 other;

Query Match 100.0%; Score 8; DB 20; Length 10;

Best Local Similarity 100.0%; Pred. No. 5.3e+04;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
Db 8 CGCGCGCG 1

RESULT 41

AAA10396
ID AAA10396 standard; DNA; 12 BP.

XX AAA10396;

XX 03-JUL-2000 (first entry)

DE Kinetic strand displacement assay duplex competition oligo, SEQ ID NO.79.

XX Nucleic acid ligand binding assay; duplex formation; stability;

XX detectable signal; kinetic strand displacement assay;

XX competition oligonucleotide; ss.

XX Synthetic.

XX WO200015848-A1.

XX 23-MAR-2000.

XX 10-SEP-1999; 99WO-US20719.

XX 11-SEP-1998; 98US-0151890.

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX Schroth GP, Bruce TW, Suh YJ;

XX WPI; 2000-271478/23.

XX Determining binding affinity of a ligand to an oligonucleotide sequence

XX in double stranded form, comprises measuring the effect of adding

XX increasing amounts of a ligand on a signal generated by two indicator

XX oligonucleotides of the duplex -

XX Example 12; Page 23; 78pp; English.

XX The invention relates to new methods of determining the binding affinity

XX of a ligand to an oligonucleotide sequence, particularly to a duplex.

XX The ligand is typically a metal ion, a small organic or inorganic

XX molecule, a protein or a multi-protein complex. The methods comprise

XX measuring the effect of adding increasing amounts of a ligand on a signal

XX generated by two indicator oligonucleotides of the duplex. In the absence

XX of ligand, conditions are such that the oligonucleotides exist primarily

XX in single-stranded form; binding of ligand to double-stranded nucleic

XX acids stabilises the duplexes, such that duplex formation is favoured.

XX One of the indicator oligonucleotides contains a first group capable of

XX producing a detectable signal, while the other indicator oligonucleotide

XX contains a second group that on hybridisation of the two indicator

XX molecules, will detectably alter the signal produced by the first group.

XX The signal may be increased or decreased on hybridisation. For example,

XX the pairs of signalling groups used could be a radioactive group and a

XX scintillant (where an increase in signal intensity indicates that

XX hybridisation has taken place) or a fluorophore and a fluorescence

XX quencher (where a reduction in signal intensity indicates that

XX hybridisation has occurred). Other methods of the invention comprise a

XX strand displacement assay, where the ability of an unlabelled

XX displacement strand to displace one of the oligonucleotides in the duplex

XX is determined in the absence and presence of ligand; and a competition

XX assay, where an unlabelled single or double-stranded competitor

XX oligonucleotide is added to the ligand-bound indicator duplex, and the

XX effect on the signal produced from the indicator duplex determined. The

XX methods are useful for determining the binding affinity of a ligand to

XX an oligonucleotide sequence. They are particularly useful for

XX determining relative binding affinities of various ligands to various

XX oligonucleotide sequences, particularly double-stranded oligonucleotide

XX sequences. The assays allow rapid and convenient determination of nucleic

XX acid binding specificities. Sequences AAA10395-A10396 represent

XX competition oligonucleotides used in a kinetic strand displacement assay

XX in an exemplification of the present invention.

XX Sequence 12 BP; 0 A; 8 C; 4 G; 0 U; 0 other;

Query Match 100.0%: Score 8; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
DB 3 CGCGCGCG 10

RESULT 42

AAA10396/c
ID AAA10396 standard: DNA; 12 BP.

XX AAA10396;

DT 03-JUL-2000 (first entry)

DE Kinetic strand displacement assay duplex competition oligo, SEQ ID NO:79.

KW Nucleic acid ligand binding assay; duplex formation; stability;

KW detectable signal; kinetic strand displacement assay;

KW competition oligonucleotide; ss.

XX Synthetic.

OS WO200015848-A1.

PN 23-MAR-2000.

PD 10-SEP-1999; 99MO-US20719.

PF 11-SEP-1998; 98US-0151890.

PR (GENE-) GENELABS TECHNOLOGIES INC.

PA Schroth GP, Bruce TW, Suh YJ;

PI WPI: 2000-271478/23.

DR Determining binding affinity of a ligand to an oligonucleotide sequence
XX in double stranded form, comprises measuring the effect of adding
XX increasing amounts of a ligand on a signal generated by two indicator
XX oligonucleotides of the duplex -

PS Example 12; Page 23; 78pp: English.

XX The invention relates to new methods of determining the binding affinity
XX of a ligand to an oligonucleotide sequence, particularly to a duplex.
XX The ligand is typically a metal ion, a small organic or inorganic
XX molecule, a protein or a multi-protein complex. The methods comprise
XX measuring the effect of adding increasing amounts of a ligand on a signal
XX generated by two indicator oligonucleotides of the duplex. In the absence
XX of ligand, conditions are such that the oligonucleotides exist primarily
XX in single-stranded form; binding of ligand to double-stranded nucleic
XX acids stabilises the duplexes, such that duplex formation is favoured.
XX One of the indicator oligonucleotides contains a first group capable of
XX producing a detectable signal, while the other indicator oligonucleotide
XX contains a second group that on hybridisation of the two indicator
XX molecules, will detectably alter the signal produced by the first group.
XX The signal may be increased or decreased on hybridisation. For example,
XX the pairs of signalling groups used could be a radioactive group and a
XX scintillant (where an increase in signal intensity indicates that
XX hybridisation has taken place) or a fluorophore and a fluorescence
XX quencher (where a reduction in signal intensity indicates that
XX hybridisation has occurred). Other methods of the invention comprise a
XX strand displacement assay, where the ability of an unlabelled
XX displacement strand to displace one of the oligonucleotides in the duplex
XX is determined in the absence and presence of ligand; and a competition
XX assay, where an unlabelled single or double-stranded competitor
XX oligonucleotide is added to the ligand-bound indicator duplex, and the
XX effect on the signal produced from the indicator duplex determined. The
XX methods are useful for determining the binding affinity of a ligand to

CC an oligonucleotide sequence. They are particularly useful for
CC determining relative binding affinities of various ligands to various
CC oligonucleotide sequences, particularly double-stranded oligonucleotide
CC sequences. The assays allow rapid and convenient determination of nucleic
CC acid binding specificities. Sequences AAA10395-A10396 represent
CC competition oligonucleotides used in a kinetic strand displacement assay
CC in an exemplification of the present invention.
XX
SQ Sequence 12 BP; 0 A; 8 C; 4 G; 0 U; 0 other;

Query Match 100.0%: Score 8; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.1e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGCGCG 8
DB 10 CGCGCGCG 3

RESULT 43

AA13767
ID AA13767 standard: DNA; 12 BP.

XX AA13767;

DT 08-MAY-2002 (first entry)

DE Simple sequence repeat, SSR, #39.

KW Simple sequence repeat; plant; ds; SSR; ryegrass; fescue; tandem repeat;

KW cereal profiling; grass profiling; seed batch purity testing.

XX Festuca arundinacea.

PN NZ509193-A.

PD 25-MAY-2001.

PE 03-JAN-2001; 2001NZ-0509193.

PF 24-DEC-1999; 99AU-0004906.

PR 04-MAY-2000; 2000AU-0007310.

PA (SAUS-) STATE SOUTH AUSTRALIA SOUTH AUSTRALIAN R.

PA (UVSC-) UNIV SOUTHERN CROSS.

PA (VICT-) STATE VICTORIA DEPT NATURAL RES & ENVIRO.

PA (UYAD-) UNIV ADELAIDE.

PA (ITMA-) INT MAIZE & WHEAT IMPROVEMENT CENT.

PI Forster JW, Jones ES;

PN WPI: 2001-512563/56.

PS Example 1; Fig 6; 72pp: English.

XX The invention relates to a substantially purified or isolated nucleic
XX acid (1) from ryegrass or fescue species including a simple sequence
XX repeat (SSR), having 2 or more tandemly repeated nucleotide core elements
XX 2-6 nucleotides in length. Also included are a nucleic acid primer
XX suitable for amplifying an SSR, identifying (M1) an SSR by preparing a
XX library of ryegrass or fescue genomic DNA enriched for SSRs and
XX identifying clones in the library containing SSRs, a library of ryegrass
XX or fescue genomic DNA enriched for SSRs prepared by the M1, selecting for
XX a gene in grass or cereal breeding by identifying an SSR that is closely
XX associated with the gene such that the SSR and the gene are
XX preferentially co-inherited, and selecting for the SSR in the
XX breeding, a method for DNA profiling grass or cereal species varieties by
XX assessing variation between SSR varieties and testing the purity of grass

or cereal seed batches by assessing variation within seed batch of an
 CC SSR. The SSRs may be used in the selection of genes in grass or cereal
 CC breeding, for profiling grass or cereal species varieties, for testing
 CC the purity of grass or cereal seed batches, and for DNA profiling to
 CC establish the distinct identity, uniformity and/or stability of a
 CC cultivar. The present sequence is a ryegrass or fescue SSR.

XX Sequence 12 BP; 0 A; 5 C; 6 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 23; Length 12;

Best Local Similarity 100.0%; Pred. No. 5.1e+04; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Indels 0; Gaps 0;

DB 4 CGCGCCGC 11

RESULT 44
 AAS13767/c
 ID AAS13767 standard; DNA; 12 BP.

XX AAS13767;

XX 08-MAY-2002 (first entry)

XX Simple sequence repeat, SSR, #39.

XX Simple sequence repeat; plant; ds; SSR; ryegrass; fescue; tandem repeat;
 KW cereal profiling; grass profiling; seed batch purity testing.

XX Festuca arundinacea.

XX NZ509193-A.

XX 25-MAY-2001.

XX 03-JAN-2001; 2001NZ-0509193.

XX 24-DEC-1999; 99AU-0004906.

XX 04-MAY-2000; 2000AU-0007310.

XX (SNUS-) STATE SOUTH AUSTRALIA SOUTH AUSTRALIAN R.

XX (VISC-) UNIV SOUTHERN CROSS.

XX (VICT-) STATE VICTORIA DEPT NATURAL RES & ENVIRO.

XX (UTAD-) UNIV ADELAIDE.

XX (ITMA-) INT MAIZE & WHEAT IMPROVEMENT CENT.

XX Forster JW, Jones ES;

XX WPI: 2001-512563/56.

XX New simple sequence repeats having 2 or more tandemly repeated
 PT nucleotide core elements isolated from ryegrass and fescue, useful for
 PT selecting of genes in grass or cereal breeding or profiling grass or
 PT cereal species varieties -

XX Example 1: Fig 6; 72pp; English.

XX The invention relates to a substantially purified or isolated nucleic
 CC acid (1) from ryegrass or fescue species including a simple sequence
 CC repeat (SSR), having 2 or more tandemly repeated nucleotide core elements
 CC 2-6 nucleotides in length. Also included are a nucleic acid primer
 CC suitable for amplifying an SSR, identifying (M1) an SSR by preparing a
 CC library of ryegrass or fescue genomic DNA enriched for SSRs and
 CC identifying clones in the library containing SSRs, a library of ryegrass
 CC or fescue genomic DNA enriched for SSRs prepared by the M1, selecting for
 CC a gene in grass or cereal breeding by identifying an SSR that is closely
 CC associated with the gene such that the SSR and the gene are
 CC preferentially co-inherited, and selecting for the SSR in the
 CC breeding, a method for DNA profiling grass or cereal species varieties by
 CC assessing variation between SSR varieties and testing the purity of grass
 CC or cereal seed batches by assessing variation within seed batch of an

CC SSR. The SSRs may be used in the selection of genes in grass or cereal
 CC breeding, for profiling grass or cereal species varieties, for testing
 CC the purity of grass or cereal seed batches, and for DNA profiling to
 CC establish the distinct identity, uniformity and/or stability of a
 CC cultivar. The present sequence is a ryegrass or fescue SSR.

XX Sequence 12 BP; 0 A; 5 C; 6 G; 1 T; 0 other;

Query Match 100.0%; Score 8; DB 23; Length 12;

Best Local Similarity 100.0%; Pred. No. 5.1e+04; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Indels 0; Gaps 0;

DB 11 CGCGCCGC 4

RESULT 45
 ABH86501
 ID ABH86501 standard; DNA; 12 BP.

XX ABH86501;

XX 22-FEB-2002 (first entry)

XX Oligonucleotide primer SEQ ID NO 286494 for detecting SNP TSC0012735.

XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;

XX peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;

XX central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX Homo sapiens.

XX WO200177384-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-IB00713.

XX 07-APR-2000; 2000DE-1019173.

XX (EPIC-) EPIDEMIOLOGICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI: 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is
 PT designed to detect single nucleotide polymorphisms and cytosine
 PT methylation status -

XX Claim 1; SEQ ID 286494; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic
 CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
 CC and cytosine methylation status in chemically pretreated genomic DNA. The
 CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
 CC range of diseases including immune system, gastrointestinal, respiratory,
 CC central nervous system, cardiovascular and metabolic disorders. The
 CC oligomers are also used for detecting cell type differentiation.

CC AB000010-AB099989, ABF00010-ABF99989, ABH00010-ABH99989 and
 CC AB100010-AB182073 represent the oligomers described in the invention.

CC NOTE: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format from WIPO at
 CC ftp.wipo.int/pub/published_pcr_sequences.

XX Sequence 12 BP; 0 A; 4 C; 6 G; 2 T; 0 other;

Query Match 100.0%; Score 8; DB 23; Length 12;

Best Local Similarity 100.0%; Pred. No. 5.1e+04; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 CGCGCCGC 8

Mon Mar 17 08:44:04 2003

cgcgcgcg.rng

Page 23

Db 1 |||||
1 CGCGCGCG 8

Search completed: March 14, 2003, 03:41:37
Job time : 198 secs

